APPLIED METHODS OF STATISTICAL ANALYSIS.

STATISTICAL COMPUTATION AND SIMULATION

PROCEEDINGS

OF THE INTERNATIONAL WORKSHOP

18-20 September 2019

Novosibirsk

2019

UDC 519.22(063)
A 67

E d i t o r s:

Prof. Boris Lemeshko, Prof. Mikhail Nikulin,
Prof. Narayanaswamy Balakrishnan

A 67 **Applied Methods of Statistical Analysis. Statistical Computation and Simulation** - AMSA'2019, Novosibirsk, Russia, 18-20 September, 2019: Proceedings of the International Workshop. - Novosibirsk: NSTU publisher, 2019. - 574 pp.
ISSN 2313-870X

ISSN 2313-870X **UDC 519.22(063)**

Applied Methods of Statistical Analysis. Statistical Computation and Simulation

Scientific Program Committee:

A. Vostretsov	Novosibirsk State Technical University, Russia
B. Lemeshko	Novosibirsk State Technical University, Russia
N. Balakrishnan	McMaster University, Canada
A. Medvedev	Siberian Federal University, Russia
G. Koshkin	Tomsk State University, Russia
A. Antonov,	Institute of Nuclear Power Engineering, Russia
E. Chimitova,	Novosibirsk State Technical University, Russia
Yu. Dmitriev,	Tomsk State University, Russia
M. Krnjajić,	National University of Ireland, Ireland
H. Liero,	University of Potsdam, Germany
N. Limnios,	Université de Technologie de Compiègne, France
G. Mikhailov,	Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Russia
V. Melas,	St. Petersburg State University, Russia
V. Ogorodnikov,	Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Russia
B. Ryabko,	Siberian State University of Telecommunications and Information Sciences, Russia
V. Rykov,	Institute of Computational Technologies SB RAS, Russia
V. Solev,	St.Petersburg Department of Steklov Mathematical Institute RAS, Russia
F. Tarasenko,	Tomsk State University, Russia
V. Timofeev	Novosibirsk State Technical University, Russia

Local Organizing Committee:

Ekaterina Chimitova, Evgenia Osintseva, Mariia Semenova

Preface

The Fifth International Workshop "Applied Methods of Statistical Analysis. Statistical Computation and Simulation" – AMSA'2019 is organized by Novosibirsk State Technical University.

The first two Workshops AMSA'2011 and AMSA'2013, as well as AMSA'2019, took place in Novosibirsk. AMSA'2015 was held in the resort Belokurikha located at the foothills of Altai. AMSA'2017, organized together with Siberian State University of Science and Technologies called after academician M.F. Reshetnev, took place in Krasnoyarsk.

The First Workshop "Applied Methods of Statistical Analysis" AMSA'2011 was focused on Simulations and Statistical Inference, AMSA'2013 – on Applications in Survival Analysis, Reliability and Quality Control, AMSA'2015 – on Nonparametric Approach and AMSA'2017 – on Nonparametric Methods in Cybernetics and System Analysis.

The Workshop AMSA'2019 was mainly oriented to the discussion of problems of Statistical Computation and Simulation, which are crucial for the development of methods of applied mathematical statistics and their effective application in practice.

The Workshop proceedings would certainly be interesting and useful for specialists, who use statistical methods for data analysis in various applied problems arising from engineering, biology, medicine, quality control, social sciences, economics and business. The Proceedings of International Workshop "Applied Methods of Statistical Analysis" are indexed in Scopus starting with 2017 materials.

The organization of the Fifth International Workshop "Applied Methods of Statistical Analysis. Statistical Computation and Simulation" – AMSA'2019 was supported by the Russian Ministry of Education and Science (project 1.1009.2017/4.6).

Prof. Boris Lemeshko

Contents

Actuarial risk theory: becoming in Russia, main problems, and development of concepts

Yu. D. Grigoriev
Saint Petersburg Electrotechnical University (LETI), Saint Petersburg, Russia
e-mail: yuri_grigoriev@mail.ru

Abstract

The subject of actuarial mathematics and its formation in Russia are considered. The definition of risk is given and the nature of risk functionals is discussed. A differentiation between decisive functionals and risk measures is indicated. Examples of order relations between risks, problems of risk management in reinsurance, are provided.

The areas adjacent with the modern risk theory including experimental design, navigating problems of vessel's place definition and problems of the reliability theory are listed. Examples of specific risk measures including measures of expected utility, measures of disturbed probability and quantile risk measures are given.

Keywords: actuarial mathematics, risk measure, coherent and comonotonic risks, expected utility, disturbed and quantile measures.

Introduction

The actuarial risk theory takes up an intermediate position between economics which dictates it the purposes and problems, and applied mathematics from which it draws methods of their decision. The problems and methods applied in the insurance companies for practical calculations are concerned with contemporary risk theory, which also includes theoretical designs, which allow to sound the actuarial methods from omnibus approach, for example, in utility theory, or ordering risk theory.

Nowadays many university courses of actuarial mathematics includes some results on net premium principles calculation, reinsurance models, properties of underlying functionals, etc. In the report the brief information on history of actuarial mathematics in Russia is presented and some specific examples of actuarial problems are provided.

1 Becoming of Actuarial Mathematics in Russia

The actuarial mathematics and actuarial education have a long history. Suffice it to mention names of E. Halley (1656-1742) and A. de Moivre (1667-1754), which in the modern terminology were the first actuaries, the foundation in 19th century the Institute of Actuaries (United Kingdom, Oxford, 1848) and Faculty of Actuaries (Glasgow, 1856), the subsequent including of actuarial calculations and methods in the higher education system, the carrying out of the International Actuarial Congresses on the

regular basis as from 1895, on which the basic directions of development actuarial science and education all over the world are defined.

In particular one of eminent actuary was famous Swedish mathematician and statistician H. Cramér, since 1918 he had been worked for insurance company. He was the founder of the first-ever chair of actuarial mathematics at Stockholm University (1929).

In our country the actuarial science also had its own traditions (Grigoriev [18]). Suffice it to mention a name of Russian mathematician S. E. Savich (1864-1946), who was a vice-president of four earliest Actuarial Congresses (1895, Brussels; 1898, London; 1900, Paris; 1903, New York), an initiator of VIII Actuarial Congress in Saint Petersburg (1915), canceled due to World War I, Professor and Head of higher mathematics chair of Saint Petersburg Electrotechnical Institute (nowadays Saint Petersburg State Electrotechnical University («LETI»)).

His contribution to development of actuarial sciences was so considerable, and it is confirmed by republishing in 2003 his book «The Elementary Theory of Life Insurance and Work Capacity» (1900) (Publishing house «Janus-K», Moscow), which is still topical up to now.

Insurance as one of institution of finance management was missing in Soviet Union. Actuarial profession, which subject is studying of economy of the finance, the theory of life insurance and the risk theory, as providing it, did not exist too. It is significant that inevitable mentions of the risk theory in foreign literature in translations (Prabhu [40, p. 215, section 5.5], [41]), as a rule, stayed out of reader's sight as it was hidden in notions of problems inventory systems and the queues theory. For instance a subtitle of the book by Prabhu [41] «Queues, risk insurance, dams» was not entered in a cover and was not included in the book.

The beginning of the refoundation of the actuarial direction in education and the appearance of a new type of insurance business in our country starts with the Diploma-Courses (Grigoriev [16]), organized by the initiative of Institute of Actuaries (United Kingdom) in 1994-1998 in different cities of Russia (Kemerovo, Novosibirsk, Moscow, St. Petersburg, Ufa) and in some former USSR republics (Belarus, Latvia). The author of this report was directly involved in organizing and conducting these courses in Novosibirsk (1996/97).

The questions of mathematical insurance theory in many domestic scientific editions are considered now. Russian actuarial science gradually gains in strength. An accruing growth both journal and book publications tells about it, the appearance of actuarial and the financial mathematics chairs at various universities, the organisation and holding of conferences on actuarial mathematics and adjacent questions, defences of master's theses (Esin[12]; Kovaleva [28], etc.), of PhD's thesis (Le Din Shon [29]; Martynova [36]) and DSci's thesis (Malinovsky [34]; Shorgin [45]). There were monographs and training manuals (Alexeev, Egorov, Ivanitsky [1]; Bojkov [6]; Bulinskaya [8]; Falin [13]; Glukhova, Zmeev, Livshits [14]; Golubin [15]; Grigoriev [18]; Kan [24]; Korolev [26]; Korolev, Bening, Shorgin [27]; Medvedev [37]; Novoselov [39]; Semenov [44]; Urazaeva [46]; Vinogradov [47]), some books have been translated (Kaas, etc. [23]; Lemaire [30], [31]; Mack [33]).

In Moscow in 1997 the International conference «Actuarial science: theory, education and applications», and in Krasnoyarsk I and II All-Russian conferences on financial and actuarial mathematics in 2002 and 2003 have been hold. In the period from 2004 to 2019 sixteen FAMEMS conferences on financial and actuarial mathematics were held, which have become International since 2009.

In St. Petersburg with support of the Russian Foundation of Fundamental Researches Annual International Science School on theme «Simulation and Analysis of Safety and Risk at Complex Systems» (MABR) by Institute of Machine Science Problem of the Russian Academy of Sciences are being held since 2001. Within the framework of this schools two sections «Technologies and reliability models in the technical Systems» and «Technologies and models of risk management in business and the finance» are hold. Numerous network conferences are also hold.

At the moment the actuarial education in Russia has not yet received organizationally complete form. Such educational programs in various institutes of higher education, including universities, are realized. It has the different levels and various forms.

Therefore it is of interest to dwell on some fundamental directions of actuarial science development from the point of view of relevant specialists' training in the system of higher education.

2 Risk and Associated Problems

Let us define a subtle distinction in risk measures and decisive functionals based on an informal notion of risk. We will consider some specific examples of risk measures and decisive functionals based on the introduced class of coherent risk measures.

2.1 Basic concepts

The basic concepts of the risk theory include the *risk*, the *order relations* between risks, the *risk measures* and their classification. On the basis of these concepts various problems of risk management are formulated.

Risk. Speaking of risk when choosing one or another decision-making strategy, in some cases risk means the probability of damage (loss) occurrence, in others cases it means the extent of damage. (Grigoriev [18, p. 141]). Depending on a situation (they are for examples, actuarial risk theory, decision-making theory, reliability theory and etc.) the risk means a random variable X, describing the extent of damage or, equivalently, its distribution function $F(x) = \mathbb{P}\{X \leq x\}$. Sometimes instead of X one could speak about a probability measure $\mathbb{P}$, that generates corresponding random variable X.

If we assume that a loss is positive, when random variable X is considered as positive, and its distribution will be concentrated in positive semiaxis $\mathbb{R}_+$. Between risks Y, Z the various partial order relations may be established.

Risk measures and decisive functionals. Not all quality functionals studying in various theories may be considered as *risk measures*. For example, *coherent* «risk measure» in practice is «decisive functional», instead of a risk measure in its pure form, as, for example, a *variation*. From the financial point of view one could say that decisive functional takes into account not only the *risk* (a probability of its occurrence), but also its *extent* (an income or a loss), and its optimization allows to make risk-weighted decisions.

Expected utility is a decisive functional, but not a coherent. Contrarily the mean $\mathbb{E}[X]$ is nominally a coherent risk measure though it fail takes risk into account as the probability of its occurrence. This measure reflects only the extent of damage, but it doesn't reflect a probability of its appearance.

Risk insurance. To make a long story short the main tasks of actuarial risk theory lie in the field of risks arising in coinsurance, namely when they are distributed between the insured and the insurer in issue of policy, in using a franchise, or in executing of various reinsurance contracts.

As a result of these actions there are various variants of risk management, because corresponding functions of a risk division are characterized by certain parameters, and decision-makers connected with corresponding utility functions on the basis of which many constructions of actuarial mathematics (inequalities of an insured and an insurer, principles of premiums awarding, risk measures, etc.) are forming.

2.2 Stochastic orders

Stochastic orders are partial orders for probability distributions. Let us consider some stochastic orders, most often used in the risk theory. Put down that $0 < \mathbb{E}[Y] < \infty$. One could say that

1. Two non-negative risks Y and Z are said to be in the *risk aversion order*, written $Y \preceq_A Z$, if and only if

$$F_Z(x) = I(x - r), \quad r > 0, \quad \mathbb{E}[Y] = r. \tag{1}$$

   Here I is a Heaviside step function, namely $I(x) = 0$ if $x < 0$, and $I(x) = 1$ if $x \geq 0$.

2. Two non-negative risks Y and Z are said to be in the *stochastic domination order*, written $Y \preceq_{ST} Z$, if and only if

$$F_Z(x) \leq F_Y(x); \tag{2}$$

3. Two non-negative risks Y and Z are said to be in the *danger order*, that is the risk Y *is less dangerous* than the risk Z, written $Y \preceq_D Z$, if $\mathbb{E}[Y] < \mathbb{E}[Z]$, and also there is a point $c \in [0, \infty)$ such that

$$F_Y(x) \leq F_Z(x) \quad at \quad x < c, \quad F_Y(x) \geq F_Z(x) \quad at \quad x > c. \tag{3}$$

4. Two non-negative risks Y and Z are said to be in the *stop-loss order*, written $Y \preceq_{ST} Z$, if

$$\forall x \in [0,\infty)\colon \int_x^\infty [1 - F_Y(t)]dt \le \int_x^\infty [1 - F_Z(t)]dt. \tag{4}$$

There are various relationships in entered order relations. In particular, relationship between orders $\preceq_{ST}$, $\preceq_D$ and $\preceq_{SL}$ establishes by following

Theorem 1. (Novoselov [39, p. 24]). *If risks Y and Z have finite means then*

$$Y \preceq_{ST} Z \Rightarrow Y \preceq_D Z \Rightarrow Y \preceq_{SL} Z. \quad \square \tag{5}$$

2.3 Coherent measures of risk

Various functionals one or another useful properties with could be implemented for the risk measurement. An interesting class of risk measures is formed with so-called *coherent measures*. They have been entered axiomatically in (Artzner [4]) and since then they are objects of intensive research from both the point of view of studying of their properties, and from their possible generalizations (Esin [10]; Martynova [36]).

Let $\mathcal{X}$ be a set of all risks on measurable space $(\Omega, \mathcal{A})$, $X, Y \in \mathcal{X}$. A coherent measure of risk is a functional $f\colon \mathcal{X} \to \mathbb{R}$ with the properties of monotony, superadditivity (subadditivity), positive uniformity, and invariance on shift transformation:

$$\begin{aligned}
X \le Y &\Rightarrow f(x) \le f(Y) \quad (f(x) \ge f(Y)), && (6)\\
f(X+Y) &\ge f(X) + f(Y), \quad (f(X+Y) \le f(X) + f(Y)), && (7)\\
f(\lambda X) &= \lambda f(x), && (8)\\
f(X+a) &\ge f(X) + a. && (9)
\end{aligned}$$

Let us notice that from (8) it follows $f(0) = 0$ that together with (6) gives $f(X) \ge 0$ for $X \ge 0$. The opposite inequalities in round brackets in (6) and (7) give an opportunity to consider the risk with an opposite sign.

Example 1. *Let $f(X) = \mathbb{E}[X]$. In this case properties* (6) – (9)*, obviously, are carried out. Hence, $\mathbb{E}[X]$ is a coherent measure of risk. Let us consider a variance $\mathbb{D}[X]$. It is easy to see that properties of uniformity* (8) *and invariance on shift transformation* (9) *are not hold. Hence a variance is not a coherent measure of risk.*

Other examples of coherent measures are the functional of so-called disturbed probability and a risk measure $CVaR$ (Hürliman [22]) *as its special case.* $\square$

2.4 Relationships with other areas

Problems similar to division of risks (though their names were different) appear in many closely related theories which have arisen long before the risk theory that has emerged clearly in the last 30-40 years.

Let us note that the actuarial science itself, as an integral part of the insurance institution, was founded much earlier, in the 18th century (E. Halley, A. de Moivre).

Closely related with the risk theory applied areas. Here we are going to list the areas of applied mathematics where the actuarial risk theory ideas are discussed:

- ☐ *Financial Tools (options, futures and other derivatives).* This is an immense research area, the list of references is unbounded.

- ☐ *Probability Theory and Stochastic Processes* (Vinogradov [47]): There are different problems: a problem about the ruining of the player (Huygens, 1657) as a problem of the risk theory. A problem of the attainment of level in classical model of collective risk. The relationship with the problems on random walk, balloting and branching processes theory.

- ☐ *Reliability Theory and Queuing Theory* (Arfvedson [2], [3]; Prabhu, 1969 [40]; 1984 [41]; Vinogradov [47]): it is a problem on ruining in the collective risk model in its connection with the queuing theory (system $M/G/1$, single-channel queueing system with Poisson input flow and arbitrary service time). The system $G/M/1$ is dual to system $M/G/1$ and consequently it is connected also with collective risk model. Notice that in Russian translation of the book title of Prabhu [41] its subtitle *«Queues, Insurance Risk and Dams»* was omitted. Obviously it shows an applicability of the inventory theory to called areas.

  The other case is the comparison of tail heaviness for various reliability functions (Proshan [5, p. 324]) and it's relationship to the problems of the risk ordering in actuarial risk theories (Kaas, etc. [23, p. 278]).

- ☐ *Systems of Maintenance Service with Periodic Inspection* (Livshits, Golichenko [32]): the problem of determination of service optimum interval T at two types of costs.

- ☐ *Inventory Systems* (Hadley, Whitin [21, p. 360]): problems of the newspapers seller (1888) and the problem about Christmas tree.

- ☐ *Experimental Design* (Grigoriev [20]): A-optimality criterion.

- ☐ *Navigation* (Kondrashikhin [25]): a problem of determining the vessel's location.

- ☐ *Control Theory* (Grigoriev, Le Din Shon [19]): Ruin probability in the collective risk model of Cramér and Lundberg at excess reinsurance. The equation of Jacobi–Hamilton–Bellman.

Let's give an example from experimental design for characterizing the similarity of statements of problems in this area with risk theory problems.

Regression experimental design. The fundamental property of every optimality criterion of experimental design ξ is *the ordering*, that is induced on the closed

cone $\mathcal{M}$ of nonnegative definite information matrixes $M \in \mathcal{M} \subset R^{m\times m}$, where m is a number of estimated parameters.

Let us consider a partial order relation $\geq$ on $\mathcal{M}$, defined by next condition:

$$A \geq B \Leftrightarrow A - B \geq 0 \Leftrightarrow A - B \in \mathcal{M}.$$

Such partial order relation is called *Loewner ordering* (Grigoriev [20, p. 95]).

The functional $\varphi\colon \mathcal{M} \to R^1$ is called *isotonic* if it keeps a Loewner ordering, namely If $A \geq B \Rightarrow \varphi(A) \geq \varphi(B)$, and it is called *antiisotonic* otherwise. The functional is called *monotonic*, if it is isotonic or antiisotonic.

One could consider such *matrix functions* as optimality criteria in experimental design, they are equipped with a partial order in their definition domain. Let us formulate some usual conditions imposed on such functions (Grigoriev [20, p. 95]; Pukelsheim [43]):

1°. We will say that information matrix C is not worse than matrix D relative to the criterion φ, if $\varphi(C) \geq \varphi(D)$. On that understanding of information matrixes it is reasonable to consider that the chosen criterion should be isotonic relative to the Loewner ordering:

$$\forall C, D \in \mathcal{M}\colon C \geq D \Rightarrow \varphi(C) \geq \varphi(D).$$

2°. The second property usually imposed on reasonable criterion, is *concavity* (*convexity*),

$$\varphi[(1-\alpha)C + \alpha D] \geq (1-\alpha)\varphi(C) + \alpha\varphi(D), \quad \alpha \in (0,1), \quad C, D \geq 0.$$

In other words, information cannot be increased by interpolation, otherwise the situation

$$\varphi((1-\alpha)C + \alpha D) < (1-\alpha)\varphi(C) + \alpha\varphi(D),$$

will occur. Rather than carrying out the experiment belonging to $(1-\alpha)C + \alpha D$, we achieve more information through interpolation of the two experiments associated with C and D. This is absurd. It is possible to show that concavity (convexity) and superadditivity (subadditivity) are equivalent (Pukelsheim [43, p. 115]).

3°. The third desirable property of criterion is *positive uniformity*:

$$\varphi(\lambda C) = \lambda\varphi(C), \quad \lambda > 0, \quad C \geq 0.$$

A function $\varphi\colon \mathcal{M} \to R^1$ is called *information*, if it isotonic, concave (convex) and positively homogeneous. Comparing (6) – (8) from 1° – 3°, one could conclude that coherent measures of risk form a subset of information functions set. In particular, D-, A- and E-optimality criteria are the most-used information functions in experimental design.

Let us deal on A-criterion $\varphi\colon trD(\xi) \to \min$. Such functional is monotonic (antiisotonic) (Grigoriev [20, p. 27, the theorem 1.14]) and convex on a set of information matrixes (Grigoriev [20, p. 95]):

$$tr[M^{-1}(\xi)] \leq (1-\alpha)tr[M^{-1}(\xi_1)] + \alpha tr[M^{-1}(\xi_2)], \quad 0 < \alpha < 1,$$

where $M(\xi) = (1-\alpha)M(\xi_1) + \alpha M(\xi_2)$, namely this functional is *subadditive*. It is easy show that $m^{-1}tr[\lambda D(\xi)] = (\lambda/m)tr[D(\xi)]$. Thus, the criterion of A-optimality is positively homogeneous. Moreover such criterion also has another property:

4°. A-optimality criterion is invariant relative to the shift a:

$$m^{-1}tr[D(\xi) + aI] = m^{-1}tr[D(\xi)] + a.$$

So when comparing 1° – 4° and (6) – (9), we can conclude that A-optimality criterion is a coherent measure of risk.

Property of 4° is specific to the problems connected with financial and insurance risks. The last requirement is not compulsory in a lot of technical applications, for instance, in experimental design, reliability theory, navigation and so on.

2.5 Comonotonic risks

Comonotonic risks play an important role in reinsurance problems in which a risk X is split on insurer risk $Y = g(X)$ and reinsurer risk $Z = X - g(X)$.

Let us consider as an example the stop-loss reinsurance with two retention levels. To do this we will formulate next theorem as a corollary of crossing condition belonging to Karlin, Novikoff, Stoyan, Taylor (KNST-condition):

Theorem 2. (Grigoriev [18, p. 145]; Hürliman [22]). (two-level stop-loss contract). *Let the following conditions are take place:*

1) *splitting $X = Y + Z$ of risk X is carried out according to Fig. 1 and Fig. 2:*

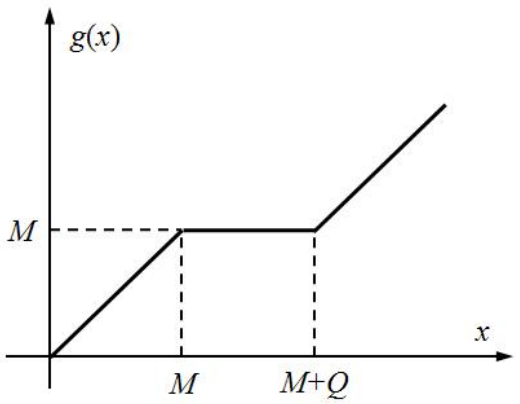


Figure 1: Insurer risk $Y = g(X)$.

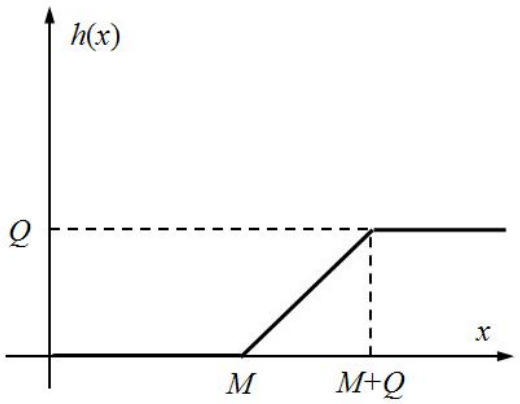


Figure 2: Reinsurer risk $Z = h(X)$.

2) *stop-loss transformations*

$$\pi_Y(x) = \int_x^\infty [1 - F_Y(t)]dt, \quad \pi_Z(x) = \int_x^\infty [1 - F_Z(t)]dt \tag{10}$$

satisfy the inequality

$$\pi_Z(M) \le \pi_Y(M).$$

Then $Z \preceq_{SL} Y$.

As it follows from the Theorem 2 (see Fig. 3 and Fig. 4) at $Q < M$ for all $x \in [0, \infty)$ an inequality $F_Y(x) \leq F_Z(x)$ is hold, namely the relation $Z \preceq_{ST} Y$ takes place.

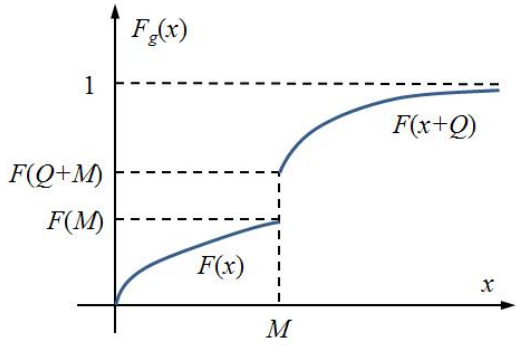


Figure 3: Two-level stop-loss contract: distribution function F_Y of an insurer.

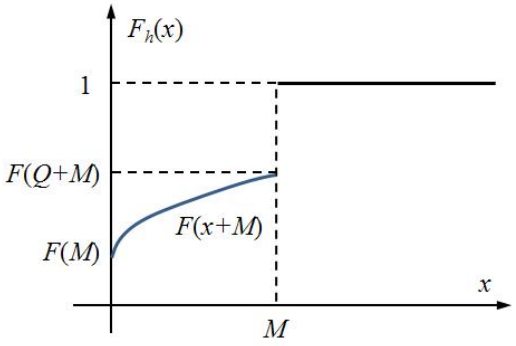


Figure 4: Two-level stop-loss contract: distribution function F_Z of a reinsurer.

This statement and Theorem 1 follow us to $Z \preceq_{SL} Y$. Thus the most interesting case is when $Q > M$, because not for any positive difference $Q - M$ the relation $Z \preceq_{ST} Y$ is realized.

3 Some specific measures of risk

Let's consider three specific risk measures as an example. First of them is not coherent measure of expected utility, and two others are coherent measures of disturbed probability. In summary we present a little quantile risk measures, among which the most important is the coherent measure *CVaR*.

3.1 Expected utility measure

The definition of utility function $u(x)$ which is a core of Neumann-Morgenstern economic theory, is given in many guides. This utility theory investigates preferences on a set of lotteries. Its basic properties are positive semidefinity of *limiting utility* $u' \geq 0$, and belonging of $u(x)$ to one of two classes $u'' \geq 0$ or $u'' \leq 0$.

Model of expected utility. The model of such type could explain the existence of insurance institute. In this model the insurer is the person, who is *not inclined to risk* and making reasonable decisions. By Jensen's inequality he is ready to pay for own financial safety more than expected value of its losses. The mechanism of decision-making under conditions of uncertainty consists not in comparison of expected payments realized as result of decisions, but in comparison to *expected utilities* these payments.

Theorem 3. (Jensen's inequality). *If $v(x)$ is a convex function and Y is a random variable, then*

$$\mathbb{E}[v(Y)] \geq v([\mathbb{E}(Y)]). \tag{11}$$

The equality in (11) *is holds if and only if v is linear on a set of concentration of a random variable Y or when $\mathbb{D}[Y] = 0$.* □

Let w be a capital of a decision-making person (DMP). From Jensen's inequality (3) it follows that for concave utility function we could obtain

$$\mathbb{E}[u(w - X)] \leq u(\mathbb{E}[w - X]) = u(w - \mathbb{E}[X]). \tag{12}$$

Therefore decision-makers with decreasing function $u''(x)$ fairly are called *as not inclined to risk*: they prefer determined (not random) $\mathbb{E}[X]$, but not random payment.

Let us briefly consider a definition, properties and an example of usage of *expected utility measures* in a problem of comparison of two risks as two lotteries (Urazaeva [46]). As it was mentioned before, the expected utility measure is a decisive functional, instead of a risk measure in its pure form.

Risk aversion ratio. The dependence of utility function curvature from «risk aversion force» allowed proposing a relative indicator *risk aversion ratio*. If the first and second derivatives of utility function are known we could get answers to following questions:

- □ What class does decision-maker belong to? Is he a riskophobe (he is not inclined to risk), a riskophile (he is inclined to risk), or a neutral person (he is indifferent to risk)?
- □ How much strong he does (or does not) accept the risk?

Based on this general knowledge about the utility function, one can get to the following definition of the risk aversion ratio. It is called *Arrow-Pratt ratio* and is defined by the formula (Malykhin [35, p. 15]; Pratt [42]):

$$r_{AP}(x) = -\frac{u''(x)}{u'(x)}, \tag{13}$$

where x is the size of decision-maker capital. If at a given level of capital x one has $r_{AP}(x) > 0$, then a case of risk aversion is taking place. Otherwise, if $r_{AP}(x) < 0$, we have a case of an inclination of decision-maker to risk.

Let X is a risk with distribution function $F(x)$, $u(x)$ is the utility function of some person. The measure of expected utility is defined as

$$\mu_u(X) := \mathbb{E}[u(X)] = \int_0^\infty u(x)dF(x). \tag{14}$$

It is said that a risk X'' is preferable to a risk X' ($X'' \succ X'$), if $\mu_u(X'') > \mu_u(X')$.

Risk as a lottery. Let us consider two risks (Urazaeva [46]):

$$X' = \begin{pmatrix} x_1 = x_0 + a, & x_2 = x_0 - b \\ p = \dfrac{b}{a+b}, & 1-p = \dfrac{a}{a+b} \end{pmatrix}, \quad X'' = \begin{pmatrix} x_1 = x_0 - a, & x_2 = x_0 + b \\ p = \dfrac{b}{a+b}, & 1-p = \dfrac{a}{a+b} \end{pmatrix}. \tag{15}$$

It is easy to verify that the following equalities are hold:

$$\mathbb{E}[X'] = \mathbb{E}[X''] = x_0, \quad \mathbb{D}[X'] = \mathbb{D}[X''] = ab.$$

Therefore it is not evident which lottery will be preferred by a person. To make it clear some additional preferences are necessary.

Example 2. (Urazaeva [46]). *Let be $x_0 = 100$, $a = 1$, $b = 99$, $p = 0.99$. Thus,*

$$\mathbb{P}\{X' = 101\} = 0.99, \quad \mathbb{P}\{X' = 1\} = 0.01, \quad \mathbb{P}\{X'' = 99\} = 0.99, \quad \mathbb{P}\{X'' = 199\} = 0.01.$$

Obviously despite the equalities of means and variances a significant number of individuals from two lotteries will prefer the $X^{\,prime\ prime}$ lottery, considering it more useful, *namely $X'' \succ X'$. But how to formalize this preference?* □

One way to develop preferences in an alternative situation is using of a *discriminating function*. Let's put

$$\Phi_\alpha(X', X'') = \mu_u(X') - \mu_u(X'').$$

If for the given utility function $u(x)$ appeared that $\Phi_\alpha(X', X'') < 0$ then one could put down $X'' \succ X'$ and on the contrary.

Theorem 4. (Urazaeva [46]). *Let the following conditions are satisfied:*
1. X', X'' are risks with the initial data presented by Example 2;
2. $u(x) = x^\alpha$, $x, \alpha \geq 0$ is utility function of the individual.
Then the following statements take place:
1. if $\alpha \in (0, 1)$ then a person is not inclined to risk and prefers a lottery X'';
2. if $\alpha \in (1, 2)$, a person inclined to risk and prefers a lottery X';
3. if $\alpha \in (2, \infty)$, a person is not inclined to risk and prefers a lottery X''.
These preferences are reflected in the discriminating function $\Phi_\mu(X', X'')$ behavior as it is depicted in Fig. 5 and Fig. 6. The minimum and maximum values of $\Phi_\mu(X', X'')$ are reached at points $\alpha_1 = 0.7668$ and $\alpha_2 = 1.8350$ accordingly. □

One can find an example in (Urazaeva cite Urazaeva2013) when for a person who is not inclined to risk the choice of risk can also be ambiguous, namely: in a case of utility function of a certain type he may prefer X'.

3.2 Disturbed probability measure

The disturbed probability measure is a generalization of the mean which is used in the most simple cases as a risk measure. It is effectively calculated in a case of discrete risk and it is investigated in details by Esin [10] – [12] and Martynova [36].

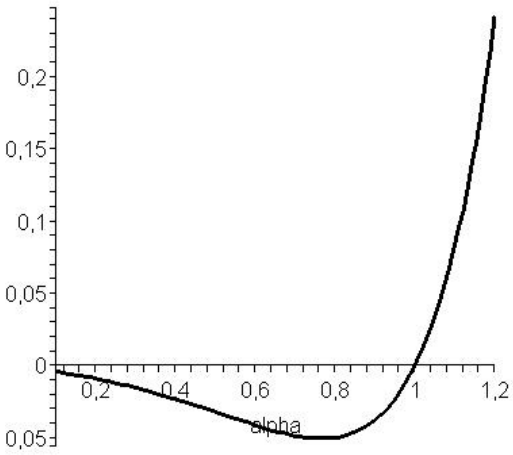


Figure 5: Measure of expected utility. (Urazaeva [46]): discriminating function $\Phi_\alpha(X', X'')$, $\alpha < 2$

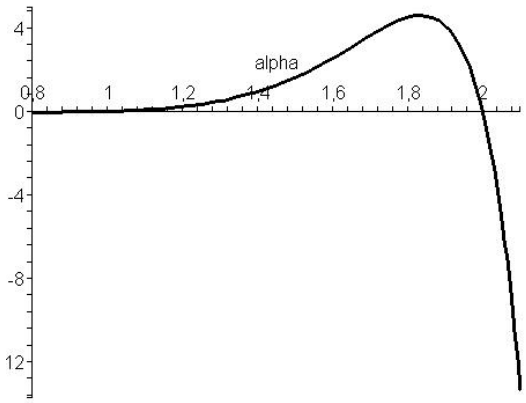


Figure 6: Measure of expected utility (Urazaeva [46]): discriminating function $\Phi_\mu(X', X'')$, $\alpha \geq 2$

Definition and calculation in a discrete case. One of the main goals of the risk theory is to construct a risk measure, which is monotonous in regard with preferences on a set of probabilistic distributions (Novoselov [38]). In (Wang [48], Young [50]) the class of risk measures named as measures of disturbed probability has been introduced, and some properties of elements of this class are investigated. The measure of disturbed probability was originally intended for calculating the insurance premium, but it can also be used in a wider class of problems, including portfolio analysis.

Let's denote $\mathcal{X}$ as a set of all real random variables and $\mathcal{X}_+$ as a set of non-negative random variables:

$$\mathcal{X}_+ = \{X \in \mathcal{X} \colon \mathbb{P}\{X \geq 0\} = 1\}.$$

And also we introduce some special notations for sets of random variables with a finite means:

$$\widetilde{\mathcal{X}} = \{X \in \mathcal{X} \colon \mathbb{E}|X| < \infty\}, \quad \widetilde{\mathcal{X}}_+ = \{X \in \mathcal{X}_+ \colon \mathbb{E}|X| < \infty\}.$$

Let further $F(x) = \mathbb{P}\{X \leq x\}$, $x \in \mathbb{R}$ is a distribution function of random variable X, and $S(x) = 1 - F(x)$ is its additional distribution function (*reliability function* in reliability theory, *survival function* in life insurance).

Let $g \colon [0,1] \to [0,1]$ is not decreasing function, and $g(0) = 0$, $g(1) = 1$. We will denote a class of all such functions as $\mathcal{G}$. It is easy to notice that for every $g \in \mathcal{G}$ corresponds to the dual function $\widetilde{g} \in \mathcal{G}$, defined by equality (Novoselov [38]):

$$\widetilde{g} = 1 - g(1 - x), \quad x \in [0,1]. \tag{16}$$

It is also evident that $\widetilde{\widetilde{g}} = g$.

In (Wong [48]) the risk measure

$$\pi_g(X) = \int_0^\infty g(S(t))dt, \quad X \in \mathcal{X}_+, \tag{17}$$

is introduced and in (Young [50]) its modification (17) for distribution of X on all real axis $\mathbb{R}$ is suggested in the form:

$$\pi_g(X) = \int_{-\infty}^0 g(S(t)-1)dt + \int_0^\infty g(S(t))dt, \quad X \in \mathcal{X}. \tag{18}$$

Risk measures (17) and (18) depend only on distribution $S(x)$ of risk X. Because it is usually $X \geq 0$, then the expression (17) is more often used as disturbed probability measure.

For discrete risk $X = \{(x_i, p_i)_{i-1}^n\}$ the measure (17) is of the form (Novoselov [39, p. 43]):

$$\pi_g(F) = \sum_{s=1}^n g\Big(\sum_{k=s}^n p_k\Big)(x_s - x_{s+1}), \quad x_0 = 0. \tag{19}$$

Example 3. Disturbed probability measures. *Let's consider the same problem, as in an Example 2. Then using the same risks X' and X'', disturbing function $g(x) = x^\alpha$, $x \in [0,1]$, $\alpha \geq 0$, discriminating function $\Phi_g(X', X'')$, in according to* (19) *we obtain:*

$$\pi_g(X') = (x_0 + a) - g(1-p)(a+b), \quad \pi_g(X'') = (x_0 - a) + (a+b)g(1-p).$$

From here it follows that

$$\frac{1}{2}\Phi_g(X', X'') = a - (a+b)(1-p)^\alpha, \quad p = b/(a+b).$$

This result is shown in Fig. 7 and one could conclude that at $\alpha \in [0,1]$ a person prefers a risk X'', and at $\alpha > 1$ he prefers a risk X'. If now we turn attention to dual function $\widetilde{g} = 1 - (1-x)^\alpha$ then we obtain

$$\frac{1}{2}\Phi_{\widetilde{g}}(X', X'') = a - (a+b)(1-p^\alpha), \quad p = b/(a+b).$$

The result is shown in Fig. 8, and it follows that now the preferences of a person have been reversed. □

Properties of a disturbed probability measure. The disturbed probability measure plays an important role in the risk theory since with an appropriate choice of function g leads to some interesting functionals from the points of view of risk theory and applications.

Let's consider some properties of functional $\pi_g(F)$.

Theorem 5. (Grigoriev [18, p. 156], Novoselov [39, p. 34-41]). *Next most important statements concerning a risk measure π take place:*

1. π is a coherent measure of risk;

2. π is an increasing functional concerning an order relation $\preceq_A$ iff $g(x) \leq x$, $x \in [0,1]$;

3. π is an increasing functional concerning an order relation $\preceq_{ST}$;

4. π is an increasing functional concerning an order relations $\preceq_D$ and $\preceq_{SL}$ iff function g is concave;

5. π is a convex functional on value iff the function g is concave. □

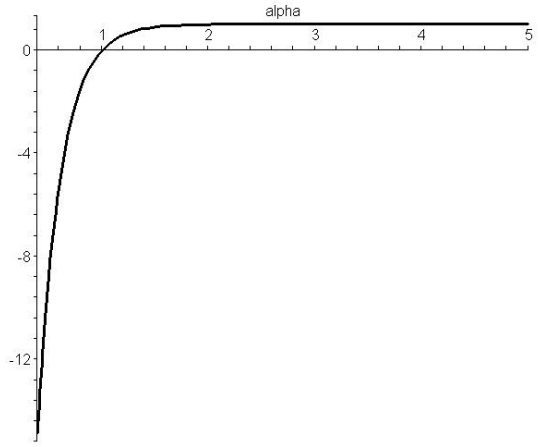


Figure 7: Measure of disturbed probability, $g(x) = x^{\alpha}$, $x \in [0,1]$: discriminating function $\Phi_g(X', X'')$

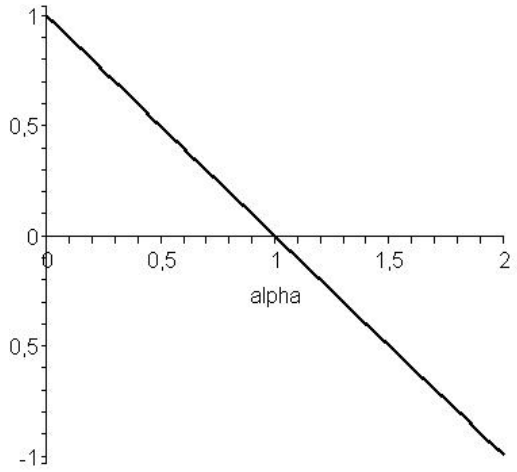


Figure 8: Measure of disturbed probability, $\widetilde{g}(x) = 1 - (1-x)^{\alpha}$, $x \in [0,1]$: discriminating function $\Phi_{\widetilde{g}}(X', X'')$

3.3 Quantile risk measures

The most in demand of financial applications is a measure $CVaR$. In practice simpler, but less reliable, measure VaR is used along with it.

Let's note another two measures connected with them, $CTER$ and ESF.

Measure *VaR*. As one of the quantile risk measures in financial and actuarial mathematics, there is a functional VaR (Value-at-Risk), which, due to the simplicity of definition, and also due to various regularity properties, is one of the most popular measures for financiers, despite the existence many other quantile risk measures.

Let $F_X(x)$ be distribution function of risk X. The value of

$$VaR_p(X) = \inf_{x \in R^1} \{F_X(x) > p\}, \quad p \in (0,1). \tag{20}$$

is called as a risk measure VaR of the level p of risk X. Risk measure VaR is often denoted by $F_X^{-1}(p)$. It is not decreasing and continuous at the left function of p.

Measure *CVaR*. Only one quantile risk measure with a predetermined level of p does not provide all the information regarding the thickness of the upper tail of the risk distribution function X. If policyholders want to assess risks in detail, they should also be interested in how bad it is. Therefore , along with VaR, they often use other risk measure named $CVaR$ (conditional VaR) of level p. Let's give an exact definition of this measure.

The value $CVaR$ is called *a measure of risk X of level p* if

$$CVaR_p(X) = \frac{1}{1-p}\int_p^1 VaR_q(X)dq, \quad p \in (0,1). \tag{21}$$

The following elementary result holds regarding the risk measure (21): if F_X is distribution function of risk X with a finite mean, $\mathbb{E}[X] < \infty$, then

$$\lim_{p\to 0} CVaR_p(X) = \mathbb{E}[X].$$

Measure *ESF*. *Expected Shortfall measure ESF of level p* for risk X is called functional

$$ESF_p(X) = \int_{VaR_p(X)}^{\infty} [1 - F_X(x)]dx, \quad p \in (0,1). \tag{22}$$

There is an equality connecting measures VaR, $CVaR$ and ESF (Grigoriev [18, p. 161]; Dhaene, et al. [9]):

$$CVaR_p(X) = VaR_p(X) + \frac{1}{1-p}ESF_p(X), \quad p \in (0,1). \tag{23}$$

As a rule, its use leads to simpler calculations of $CVaR$ compared to (21).

Measure *CTE*. *The Conditional Tail Expectation risk measure CTE of level p* on condition that losses exceed VaR is a functional

$$\begin{aligned} CTE_p(X) &= \mathbb{E}[X/X > VaR_p(X)] \\ &= VaR_p(X) + \frac{1}{1-p}\int_0^{\infty} [1 - F_X(x + VaR_p(X))]dx. \end{aligned} \tag{24}$$

The measure CTE is not quite independent because for continuous distributions F_X it coincides with $CVaR$, and also it admits another definition:

$$CTE_p(X) = CVaR_{F_X(VaR_p(X))}(X). \tag{25}$$

However in a discrete case these measures are different.

Example 4. *Let's consider an exponential risk X. In this case $F_X(x) = 1 - e^{-\lambda x}$, $x \geq 0$. Then according to equations* (20)– (22) *one could obtain:*

$$VaR_p(X) = -\lambda^{-1}\log(1-p), \quad CVaR_p(X) = \lambda^{-1} + VaR_p(X), \quad ESF_p(X) = \frac{1-p}{\lambda}.$$

Now it is easy to check the validity of equalities (23) *and* (25). □

In conclusion we formulate the theorem, in which relations of stop-loss order $\preceq_{ST}$ and stochastic domination $\preceq_{SL}$ between risks X and Y in terms of quantile risk measures VaR and $CVaR$ are characterized.

Theorem 6. (Grigoriev [18]; Dhaene, etc. [9]). *For any pair of risks the following statements are fair:*

1. $X \preceq_{ST} Y \Leftrightarrow VaR_p(X) \leq VaR_p(Y)$ *for all* $p \in (0,1)$.
2. $X \preceq_{SL} Y \Leftrightarrow CVaR_p(X) \leq CVaR_p(Y)$ *for all* $p \in (0,1)$. □

Conclusions

The considered issues of the development of the actuarial risk theories in Russia reflect the high interest to this area of researches from professional mathematicians, including the high school, where future experts are trained for the insurance sector of the Russian economy. It is essentially important that one of the national projects accepted by the government of the Russian Federation to realization in the coming years is the digital economy. One of its pivotal elements definitely is the development of the actuarial sphere in the field of economics and bank risk management.

Acknowledgements

Author expresses acknowledges to PhD T. A. Kustitskaya, PhD A. A. Novoselov (Krasnoyarsk, Siberian Federal University) and PhD V. Yu. Shchekoldin (Novosibirsk, NSTU) for useful discussions and computer support, lent me the support in training of this report.

References

[1] Alekseev B. V., Egorov D. V., Ivanitsky A. Yu. (2001). *Introduction to Financial and Actuarial Mathematics*. Chuvash State Univrersity, Cheboksary. (In Russian).

[2] Arfwedson G. (1954). Research in collective risk theory. I. *Skan. Aktuar.*, Vol. **37**, pp. 191-223.

[3] Arfwedson G. (1955). Research in collective risk theory. II. *Skan. Aktuar.*, Vol. **38**, pp. 53-100.

[4] Artzner Ph., Delbaen F., Eber J.-M., Heath D. (1999). Coherent measures of risk. *Mathematical Finance*, Vol. **9**, pp. 203-228.

[5] Barlow R. E., Proschan F. (1969). *Mathematical Theory of Reability*. John Wiley& Sons, Inc., New York, London, Sydney.

[6] Bojkov A. V. (2004). *Insurance and Actuarial Calculation*. Institute of System Analysis, Russian Academy of Sciences, Moscow. (In Russian).

[7] Bogojavlensky S. B. (2014). *Theoretical and Practical Aspects of Decision-making in the Conditions of Uncertainty and Risk*. The manual. St. Petersburg State Economic Univrersity, St. Petersburg. (In Russian).

[8] Bulinskaya E. V. (2001) *Theory of Risk and Reinsurance. Ordering of Risks. Part 2*. Moscow State University, Moscow. (In Russian).

[9] Dhaene J., Vanduffel S., Qihe Tang, Goovaerts M. J., Kaas R. Vyncke D. (2004). Solvency Capital, Risk measures and Aomonotonicity. *Review*. July 14.

[10] Esin R. V. (2013) Decision of an inverse problem of the risk theory under the model CVaR. *Proceed. of XII Intern. Conf. on Financial and Actuarial Mathematics and Eventology of Multidimensional Statistics* / Edt. Oleg Vorobjev. Siberian Federal University, Krasnoyarsk. (In Russian).

[11] Esin R. V. (2014). Decision of an inverse problem of the risk theory for a disturbed risk measure. *Proceed. of XIII Intern. Conf. on Financial and Actuarial Mathematics and Eventology of Multidimensional Statistics* / Edt. Oleg Vorobjev. Siberian Federal University, Krasnoyarsk. (In Russian).

[12] Esin R. V. (2014) *Decision of an Inverse Problem of the Risk Theory for Some Coherent Measures of Risk*. Thesis / Supervisor of Studies PhD T. A. Kustitskaya. Siberian Federal University, Krasnoyarsk. (In Russian).

[13] Falin G. I. (1994). *Mathematical Analysis of the Risks Insurance*. Russian Juridical Publishing House, Moscow.

[14] Glukhova E. V., Zmeev O. A., Livshits K. I. (2004). *Mathematical Models of Insurance*. Tomsk State University, Tomsk. (In Russian).

[15] Golubin A. Yu. (2003). *Mathematical Models in the Insurance Theory: Construction and Optimisation*. ANKIL, Moscow. (In Russian).

[16] Grigoriev Yu. D. (2003). Actuarial profession in higher education system: history of development and the basic problems. *Bulletin of St. Petersburg State Electrotechnical University («LETI»). Series «Computer science, control and computer technologies»*. Vol. **3**, pp. 46-52. (In Russian).

[17] Grigor'ev Yu. D. and Le Dinh Shon. (2005). Numerical optimisation of Value-at-Risk functional for stop-loss reinsurance for the Cramér-Lundberg risk model. *Modelling and Analysis of Safety and Risk in Complex Systems / Proceed. V Intern. Scien. School MA SR-2005*, June 28 July 1, St. Petersburg, Russia.

[18] Grigoriev Yu. D. (2006). *Models of Actuarial Mathematics and Risk Theory*. ETU, St. Petersburg. (In Russian).

[19] Grigor'ev Yu. D. and Le Dinh Shon. (2007). On Ruin probability minimisation under excess reinsurance. *Automation and Remote Control*, Vol. **68 (6)**, pp. 1039-1054.

[20] Grigoriev Yu. D. (2015). *Methods of Optimal Design of Experiment. Linear Models*. LAN, St. Petersburg. (In Russian).

[21] Hadley G., Whitin T. M. (1969). *Analysis of Inventory Sysmems*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

[22] Hürlimann W. (1998). On Stop-loss order and the distorsion pricing principle. *ASTIN Bulletin*. Vol. **28 (1)**. pp. 119-134.

[23] Kaas R., Goovaerts M., Dhaene J., Denuit M. (2007). *Modern Actuarial Risk Theory*. Kluwer Academic Publishers. Boston, Dordrecht, London.

[24] Kan Yu. S. (2006). *Risk Management and Introduction to Insurance*. Teaching aid. Moscow Aviation Institute. Moscow. (In Russian).

[25] Kondrashikhin V. T. (1989). *Definition of the Vessel Location*. Transport, Moscow. (In Russian).

[26] Korolev V. Yu. (2011). *Mathematical Foundation of the Risk Theory*. FIZMATLIT, Moscow. (In Russian).

[27] Korolev V. Yu., Bening V. E., Shorgin S. Ya. (2007). *Mathematical Foundation of the Risk Theory*. FIZMATLIT, Moscow. [A Series: Mathematics and the Applied Mathematics. The Textbook for High Schools]. (In Russian).

[28] Kovaleva I. O. (2003) *Mathematical Models of Excess Reinsurance*. Thesis / Supervisor of Studies DSc Yu. D. Grigoriev. Novosibirsk State University, Novosibirsk. (In Russian).

[29] Le Din Shon. (2007). *Development of Methods and Algorithms for Optimisation of risk at the Reinsurance problems*. Dissertation / Supervisor of Studies DSc Yu. D. Grigoriev. St. Petersburg State Electrotechnical University («LETI»), St. Petersburg. (In Russian).

[30] Lemaire J. (1995). *Bonus-Malus Systems in Automobile Insurance*. . Kluwer-Nijhoff Publishing.

[31] Lemaire J. (1996). *Automobile Insurance. Actuarial Mmodels*. Kluwer Academic Publishers.

[32] Livshits V. M, Golichenko V. I. (1976). *Principles of System Formation of Agricultural Maintenance Service under an Economy of Siberia*. Methodical recommendations. All-Union academy of agricultural sciences, Novosibirsk. (In Russian).

[33] Mack T. (2005). *Schadenversicherungsmathematik*. Deutsche Gesellschaft für Versicherungsmathematik. Karlsruhe.

[34] Malinovskii V. K. (2000) *Insurance Rating and Reserving: Stochastic Models and Methods of Computations*. D. Sci. Thesis, Steklov Math. Inst., Moscow. (in Russian).

[35] Malykhin V. I. (1998). *Mathematical Modelling of Economy*. Teaching aid. URAO, Moscow.

[36] Martynova T. A. (2007). *Generalized Coherent Measures of Risk and their Application to Decision-making Problems*. Dissertation. Siberian Federal University, Krasnoyarsk. (In Russian).

[37] Medvedev G. A. (2001). *Mathematical Models of Financial Risks. Part 2. Insurance Risks*. Belarus State University. Minsk. (In Russian).

[38] Novoselov A. A. (2000). *Selected Lectures on Modelling of Financial Risks*. Internet Resource. Siberian Federal University, Krasnoyarsk. (In Russian).

[39] Novoselov A. A. (2001). *Mathematical Modelling of Financial Risks. Measurement theory*. Science, Novosibirsk. (In Russian).

[40] Prabhu N. U. (1969). *Queues and Inventories. A study of Their Basic Stochastic Processes*. John Wiley& Sons, Inc., New York, London, Sydney.

[41] Prabhu N. U. (1984) *Stochastic Storage Processes. Queues, Insuranse Risk and Dams*. Springer-Verlag. New York, Heidelberg, Berlin.

[42] Pratt J. W. (1964) Risk aversion in the small and in the large. *Econometrica. The Econometric Society*, Vol. **32**, pp. 122-136.

[43] Pukelsheim F. (2006) *Optimal Design of Experiments*. SIAM, Philadelphia.

[44] Semenov A. T. (2009). *Introduction to General Risk Theory*. Novosibirsk State University of Economical and Management, Novosibirsk. (In Russian).

[45] Shorgin S. Ya. (1996). *Definition of Insurance Tariffs: Stochastic Models and Estimation Methods*. D. Sci. Thesis. Moscow State University, Moscow. (In Russian).

[46] Urazaeva T. A. (2013). *Algebra of Risks: Theory and Algorithms*. Volga Region State Technological University, Yoshkar-Ola. (In Russian).

[47] Vinogradov O. P. (2019). *Elements of Risk Theory*. LENAND, Moscow [Series: Classical Textbook of the Moscow State University]. (In Russian).

[48] Wang S. S. (1996) Premium calculation by transforming the layer premium density *ASTIN Bulletin*. Vol. **26 (1)**, pp. 71-92.

[49] Yaari M. E. (1987) The dual theory of choice under risk. *Econometrica*. Vol. **55**, pp. 95-115.

[50] Young V. R. (1999) Discussion of Christofide's conjecture regarding Wang's premium principle. *ASTIN Bulletin*, Vol. **29 (2)**, pp. 191-195.

Estimation of the present values of net premiums and life annuities for the different actuarial models

Yury G. Dmitriev, Oxana V. Gubina and Gennady M. Koshkin
National Research Tomsk State University, Tomsk, Russia
e-mail: dmit@mail.tsu.ru, gov7@mail.ru, kgm@mail.tsu.ru

Abstract

The paper deals with the estimation problem of the actuarial present values of the continuous individual net premiums and connected with these characteristics life annuities. We considered the following actuarial models: the whole life insurance, n-year term life insurance, q-year deferred life insurance, and n-year endowment life insurance. We synthesize nonparametric estimators of net premiums and life annuities for these statuses. The main parts of the asymptotic mean square errors of the estimators and their limit distributions are found. The simulations show that the empirical mean square errors of estimators decrease when the sample size increases. Also, when the model distribution is changed, the nonparametric estimators are more adaptable in comparison with parametric estimators, oriented on the best results only for the given distributions.

Keywords: nonparametric estimation; life insurance; net premium; life annuity; asymptotic normality; bias; mean squared error.

Introduction

One of the main issues addressed in actuarial mathematics is to find the "right" ratio between premiums and benefits, aided calculation of net premiums intended to cover damages and giving zero average income of the insurance company. Section devoted to this area in the monograph "Actuarial Mathematics" [8], in which the calculation of net premiums was based on the use of mortality tables. Interesting results based on this approach have been prepared in papers [5, 10, 13, 16, 34, 39, 42]. Modern development of theory of insurance is strongly required the use of complex mathematical models phenomena and processes taking place in this area. Note the results obtained in this direction in the papers [1, 4, 6, 18, 19, 22, 37]. Alternative solution is to build estimators of net premium functionals on the base of information containing in a sample of individuals' lifetimes. Here we develop this idea embedded in the articles [20], [25]-[33]. The second part of the paper deals with the study of life annuities estimators.

1 Individual Whole Life Insurance: Net Premiums and Life Annuities

In long-term insurance the calculations of tariff rates take into account change of money value because the sum of S dollars after t years turns to the sum $S\,e^{\delta t}$ dollars, where δ is instantaneous interest rate. The whole life insurance is example of long-term insurance; in this situation the person pays p dollars to the insurance company, and the company pays b dollars to successors of the insured after his death. Though the premium p is less, than b, the company will receive the necessary sum b, since the premium is paid at the moment of the conclusion of the contract, and the payment is done great later. We will use designations of actuarial mathematics later on. Let random variable X denote the future lifetime, x be the age of the insured at the moment of policy issue, $T(x) = X - x$ denote the residual time of life. In time $T(x)$ premium, p, will turn in the sum, $p\,e^{\delta T(x)}$, and in this case the income of the company will be equal to

$$p\,e^{\delta T(x)} - b.$$

To have the required sum b dollars at the moment of client death, the insurance company must receive $b\,e^{-\delta T(x)}$ dollars at the time of policy issue. In economic terms, the sum $b\,e^{-\delta T(x)}$ expresses discounted value of the future insurance payment. As the above mentioned this sum is a random variable, so it is natural to take as net premium its average the symbol of the expectation. In actuarial science the benefit b is accepted as a unit payment, that is, $b = 1$, and the net premium of the whole life insurance $\overline{A}_x$ is equal to $\mathbf{E}\{e^{-\delta T(x)}\}$:

$$\overline{A}_x = \mathbf{E}\{e^{-\delta T(x)}\} = -\int_0^\infty e^{-\delta t}\, d\mathbf{P}\{T(x) > t | T(x) > 0\} =$$

$$= \frac{-\int_0^\infty e^{-\delta t}\, d\mathbf{P}\{T(x) > t \cap T(x) > 0\}}{\mathbf{P}\{T(x) > 0\}} =$$

$$= \frac{\int_0^\infty e^{-\delta t}\, I_{T(x)}(t > 0) dF_{T(x)}(t)}{S_{T(x)}(0)} = \frac{\Phi(x, \delta)}{S_{T(x)}(0)}, \qquad (1)$$

where $F_{T(x)}(t) = \mathbf{P}(T(x) \le t)$ is the distribution function of the random variable $T(x)$, $S_{T(x)}(t) = 1 - F(t) = \mathbf{P}(T(x) > t)$ is the survival function, $I_{T(x)}(t > 0) = I(T(x) > 0)$, $I(A)$ is the indicator of set A.

It is known that life annuities are closely related to the corresponding net premiums (see [4]). The idea of life annuity in accordance with ([4], p. 170) is this: from the moment t = 0 an individual once a year begins to get a certain money, which we take as the unit of money, and payments are made only for the lifetime of an individual. As the calculation of the characteristics of life annuity is based on the characteristics of the respective type of insurance, the average total cost of the present continuous

annuity is defined by the following formula (see [4], p. 184):

$$\overline{a}_x(\delta) = \frac{1 - \overline{A}_x}{\delta}, \tag{2}$$

where $\overline{A}_x$ is a net premium (the average of the present value of a single sum of money in the insurance lifetime at the age x. Let us introduce the random variable

$$z(x) = \frac{1 - e^{-\delta T(x)}}{\delta}, \quad T(x) > 0. \tag{3}$$

Then, by averaging the random variable $z(x)$ (3), we get the formula of the whole life annuity (see [8, 21, 30]):

$$\overline{a}_x(\delta) = \mathbf{E}z(x) = \frac{1}{\delta}\left(1 - \frac{\Phi(x,\delta)}{S_{T(x)}(0)}\right). \tag{4}$$

2 Collective Life Insurance

A useful abstraction in the collective life insurance is that of "*status* for which there are definitions of survival and failure" [8]. Consider m members of ages $(x_1, \ldots, x_m)$ who desire to buy an insurance policy. Denote the future lifetime of the k-th individual by $T(x_k) = X_k - x_k$. Let us put in a correspondence a status U with its future lifetime $T(U)$ and with a set of numbers $T(x_1), \ldots, T(x_m)$ [20].

In the papers [25]-[33] were considered cases of a joint-life status and a last-survivor status.

The joint-life status is denoted by $U := x_1 : \ldots : x_m$ and is considered as failed upon the first death, i.e.,

$$T(U) = \min\left(T(x_1), \ldots, T(x_m)\right).$$

It is evident that the probability

$$\mathbf{P}\{T(U) > t\} = \mathbf{P}\{\min(T(x_1), \ldots, T(x_m)) > t\} = \mathbf{P}\{T(x_1) > t, \ldots, T(x_m) > t\},$$

so, when the deaths are independent, we have $\mathbf{P}\{T(U) > t\} = \prod_{i=1}^{m} \mathbf{P}\{T(x_i) > t\}$.

The last-survivor status is denoted by $U := \overline{x_1 : \ldots : x_m}$ and fails upon the last death, and exists as long as at least one member of a group is alive, i.e.,

$$T(U) = \max(T(x_1), \ldots, T(x_m)).$$

Similarly,

$$\mathbf{P}\{T(U) \le t\} = \mathbf{P}\{\max(T(x_1), \ldots, T(x_m)) \le t\} = \mathbf{P}\{T(x_1) \le t, \ldots, T(x_m) \le t\},$$

and in the case of independent deaths, we have $\mathbf{P}\{T(U) \le t\} = \prod_{i=1}^{m} \mathbf{P}\{T(x_i) \le t\}$.

We give other statuses used in practice. Consider the general k-survivor status, which is denoted

$$U := \frac{k}{x_1 : \ldots : x_m}$$

and exists as long as at least alive k among m individuals $(x_1), \ldots, (x_m)$, i.e., it is considered destroyed upon the occurrence of the $(m-k+1)$ deaths. It is understood that the joint-life status $(k = m)$ and last-survivor status $(k = 1)$ are the special cases of the k-survivor status. Also, separately consider the $[k]$-deferred survivor status

$$U := \frac{[k]}{x_1 : \ldots : x_m}$$

and there, if alive exactly k of m individuals $(x_1), \ldots, (x_m)$, i.e., it starts at the $(m-k)$-th death and lasts until the $(m-k+1)$-th death. This status is widely used in the calculation sequences payments of limited duration [20]. Note that the new statuses can be defined by compounding. A compound status is said to exist if the status is a a combination of statuses, and at least one of them is itself a status with more than one individual. Consider, for example, some compound statuses.

- **The status** $((\overline{x_1 : x_2} : \overline{x_3 : x_4})$

  This status persists if alive at least one of (x_1) and (x_2) and at least one of (x_3) and (x_4). The time-until-failure of the status $(\overline{x_1 : x_2} : \overline{x_3 : x_4})$ is

  $$T(U) = \min\{\max\{T(x_1), T(x_2)\}, \max\{T(x_3), T(x_4)\}\}.$$

- **The status** $\left(\overline{\overline{x_1 : x_2} : (x_3 : x_4)}\right)$

  Such condition persists, if alive at least two of four, namely, (x_3) and (x_4), or when only one alive, and that either (x_1), or (x_2). The time-until-failure of the status $\left(\overline{\overline{x_1 : x_2} : (x_3 : x_4)}\right)$ is

  $$T(U) = \max\{\max\{T(x_1), T(x_2)\}, \min\{T(x_3), T(x_4)\}\}.$$

- **The status** $(x_1 : x_2 : \overline{x_3 : x_4})$

  The condition persists, if alive (x_1), (x_2) and when one is alive, and it is either (x_3), or (x_4). The time-until-failure of the status $(x_1 : x_2 : \overline{x_3 : x_4})$ is

  $$T(U) = \min\{T(x_1), T(x_2), \max\{T(x_3), T(x_4)\}\}.$$

Similarly, the fracture point may be found for the combination any statuses.

3 Functionals of the Net Premiums in Collective Life Insurance

Reasoning as in the derivation of formula (1)), in the case of the m insureds, the functionals of the net premiums in collective life insurance can be written as

$$\overline{A}_{x_1:\ldots:x_m} = \frac{\int_0^\infty e^{-\delta t} I_{x_1:\ldots:x_m}(t>0) dF_{x_1:\ldots:x_m}(t)}{S_{x_1:\ldots:x_m}(0)}, \tag{5}$$

$$\overline{A}_{\overline{x_1:\ldots:x_m}} = \frac{\int_0^\infty e^{-\delta t} I_{\overline{x_1:\ldots:x_m}}(t>0) dF_{\overline{x_1:\ldots:x_m}}(t)}{S_{\overline{x_1:\ldots:x_m}}(0)},$$

where

$$F_{x_1:\ldots:x_m}(t) = \mathbf{P}(\min(T(x_1),\ldots,T(x_m)) \le t)$$

and

$$F_{\overline{x_1:\ldots:x_m}(t)} = \mathbf{P}(\max(T(x_1),\ldots,T(x_m)) \le t)$$

are the distribution functions of the random variables $\min(T(x_1),\ldots,T(x_m))$ and $\max(T(x_1),\ldots,T(x_m))$,

$$S_{x_1:\ldots:x_m}(t) = 1 - F_{x_1:\ldots:x_m}(t) = \mathbf{P}(\min(T(x_1),\ldots,T(x_m) > t))$$

and

$$S_{\overline{x_1:\ldots:x_m}}(t) = 1 - F_{\overline{x_1:\ldots:x_m}}(t) = \mathbf{P}(\max(T(x_1),\ldots,T(x_m) > t))$$

are the corresponding survival functions.

Consider the random variables $Z_i = X_i - x_i$, $i = \overline{1,m}$. We order them in ascending and obtain the order statistics $Z_{(i)}$, $i = \overline{1,m}$. Note that the survival function

$$S_{x_1:\ldots:x_m}(0) = \mathbf{P}(\min(T(x_1),\ldots,T(x_m) > 0)) = \mathbf{P}(T(x_1) > 0,\ldots,T(x_m) > 0) =$$

$$= \mathbf{P}(X_1 > x_1,\ldots,X_m > x_m) = S(x_1,\ldots,x_m).$$

Then

$$\overline{A}_{x_1:\ldots:x_m} = \frac{-\int_0^\infty e^{-\delta t} I(Z_{(1)} > 0) d[1 - \mathbf{P}\{Z_{(1)} > t\}]}{\mathbf{P}\{Z_{(1)} > 0\}} =$$

$$= \frac{-\int_0^\infty e^{-\delta t} I(\min(T(x_1),\ldots,T(x_m)) > 0) d[1 - \mathbf{P}\{\min(T(x_1),\ldots,T(x_m)) > t\}]}{\mathbf{P}\{\min(T(x_1),\ldots,T(x_m)) > 0\}} =$$

$$= \frac{-\int_0^\infty e^{-\delta t} \prod_{j=1}^m I(T(x_j) > 0) d[1 - \mathbf{P}\{(T(x_1) > t,\ldots,T(x_m) > t)\}]}{\mathbf{P}\{T(x_1) > 0,\ldots,T(x_m) > 0\}} =$$

$$= \frac{\Phi(x_1 : \ldots : x_m, \delta)}{S(x_1, \ldots, x_m)} \tag{6}$$

and

$$\overline{A}_{\overline{x_1:\ldots:x_m}} = \frac{\int_0^\infty e^{-\delta t} I(Z_{(m)} > 0) d\mathbf{P}\{Z_{(m)} \le t\}}{\mathbf{P}\{Z_{(m)} > 0\}}.$$

By analogy with formula (1), we have

$$\overline{A}_{\underset{x_1:\ldots:x_m}{^{k}}} = \frac{\int_0^\infty e^{-\delta t} I_{Z_{(m-k+1)}}(t > 0) d\mathbf{P}\{Z_{(m-k+1)} \le t\}}{\mathbf{P}\{Z_{(m-k+1)} > 0\}}, \tag{7}$$

where $I_{Z_{(m-k+1)}}(t > 0) = I(Z_{(m-k+1)} > 0)$.

In the case of the $[k]$-deferred survivor status

$$\mathbf{P}\left\{Z_{(m-k)} < t < Z_{(m-k+1)}\right\} = \mathbf{P}\left\{t < Z_{(m-k+1)}\right\} - \mathbf{P}\left\{t < Z_{(m-k)}\right\} =$$

$$= 1 - \mathbf{P}\left\{Z_{(m-k+1)} \le t\right\} - 1 + \mathbf{P}\left\{Z_{(m-k)} \le t\right\} = \mathbf{P}\left\{Z_{(m-k)} \le t\right\} - \mathbf{P}\left\{Z_{(m-k+1)} \le t\right\},$$

and the net premium is given by the formula

$$\overline{A}_{\underset{x_1:\ldots:x_m}{^{[k]}}} = \frac{\int_0^\infty e^{-\delta t} I(Z_{(m-k)} > 0) d\mathbf{P}\{Z_{(m-k)} \le t\}}{\mathbf{P}\{Z_{(m-k)} > 0\}} -$$

$$- \frac{\int_0^\infty e^{-\delta t} I(Z_{(m-k+1)} > 0) d\mathbf{P}\{Z_{(m-k+1)} \le t\}}{\mathbf{P}\{Z_{(m-k+1)} > 0\}} = \overline{A}_{\underset{x_1:\ldots:x_m}{^{k-1}}} - \overline{A}_{\underset{x_1:\ldots:x_m}{^{k}}}.$$

The functionals of net premiums for compound statuses can be written in the same way.

4 Estimators of the Net Premiums in Collective Life Insurance

Let $(Z_{11}, \ldots, Z_{m1})$, $\ldots$, $(Z_{1n}, \ldots, Z_{mn})$ be an m-dimensional random sample and $(Z_{(1)1}, \ldots, Z_{(m)1})$, $\ldots$, $(Z_{(1)n}, \ldots, Z_{(m)n})$ be corresponding ordered set.

According to (6) as the estimator of the survival function $\mathbf{P}\left\{Z_{(1)} > t\right\}$, we take $\frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{m} I\left(Z_{ji} > t\right)$. Let $\overline{\delta}(t)$ be the Dirac function. Then, the nonparametric estimator of net premium (6) is given by

$$\hat{\overline{A}}_{x_1:\ldots:x_m} = \frac{-\int_0^\infty e^{-\delta t} \prod_{j=1}^{m} I_{Z_j}(t > 0) d[1 - \mathbf{P}_n\{Z_1 > t, \ldots, Z_m > t\}]}{\mathbf{P}_n\{Z_1 > 0, \ldots, Z_m > 0\}} =$$

$$= \frac{\int_0^\infty e^{-\delta t} \prod_{j=1}^m I_{Z_j}(t>0) \left(\frac{1}{n} \sum_{i=1}^n \prod_{j=1}^m I(Z_{ji} > t) \right)' dt}{\frac{1}{n} \sum_{i=1}^n \prod_{j=1}^m I\left(Z_{ji} > 0\right)} =$$

$$= \frac{1}{\frac{1}{n} \sum_{i=1}^n \prod_{j=1}^m I\left(Z_{ji} > 0\right)} \frac{1}{n} \sum_{i=1}^n \int_0^\infty e^{-\delta t} \prod_{j=1}^m I_{Z_j}(t>0) \overline{\delta}(t - \prod_{j=1}^m Z_{ji})\, dt =$$

$$= \frac{\frac{1}{n} \sum_{i=1}^n e^{-\delta \prod_{j=1}^m Z_{ji}} \prod_{j=1}^m I(Z_{ji} > 0)}{\frac{1}{n} \sum_{i=1}^n \prod_{j=1}^m I\left(Z_{ji} > 0\right)}. \tag{8}$$

As estimators of the distribution function $\mathbf{P}\left\{Z_{(m-k+1)} \le t\right\}$ and the survival function $\mathbf{P}\left\{Z_{(m-k+1)} > 0\right\}$, we take $\frac{1}{n} \sum_{i=1}^n I\left(Z_{(m-k+1)i} < t\right)$ and $\frac{1}{n} \sum_{i=1}^n I\left(Z_{(m-k+1)i} > 0\right)$, respectively. So, the nonparametric estimator of (4) has the form

$$\hat{\overline{A}}_{\underbrace{\quad}_{x_1:\ldots:x_m}^{k}} = \frac{\int_0^\infty e^{-\delta t} I_{Z_{(m-k+1)}}(t>0) \left(\frac{1}{n} \sum_{i=1}^n I(Z_{(m-k+1)i} \le t) \right)' dt}{\mathbf{P}_n\{Z_{(m-k+1)} > 0\}} =$$

$$= \frac{1}{\mathbf{P}_n\{Z_{(m-k+1)} > 0\}} \frac{1}{n} \sum_{i=1}^n \int_0^\infty e^{-\delta t} I_{Z_{(m-k+1)}}(t>0) \overline{\delta}(t - Z_{(m-k+1)i})\, dt =$$

$$= \frac{\frac{1}{n} \sum_{i=1}^n e^{-\delta \left(Z_{(m-k+1)i}\right)} I(Z_{(m-k+1)i} > 0)}{\frac{1}{n} \sum_{i=1}^n I\left(Z_{(m-k+1)i} > 0\right))}. \tag{9}$$

In the case of the $[k]$-deferred survivor status the nonparametric plug-in estimator of the net premium can be defined in the following way:

$$\hat{\overline{A}}_{\underbrace{\quad}_{x_1:\ldots:x_m}^{[k]}} = \hat{\overline{A}}_{\underbrace{\quad}_{x_1:\ldots:x_m}^{k-1}} - \hat{\overline{A}}_{\underbrace{\quad}_{x_1:\ldots:x_m}^{k}}.$$

5 Asymptotics of the Functions of Statistics

Introduce the notation according to [12, 6]: the function $H(t) : R^s \to R^1$, where $t = t(x) = (t_1(x), \ldots, t_s(x))$ is s-dimensional bounded function; $H_j(t) = \dfrac{\partial H(t)}{\partial t_j}$,

$j = \overline{1,s}$, $\nabla H(t) = (H_1(t), \ldots, H_s(t))$; the symbol T denotes the transpose; $t_n = (t_{1n}, \ldots, t_{sn})$ is s-dimensional statistic, $t_{jn} = t_{jn}(x) = t_{jn}(x, X_1, \ldots, X_n)$, $j = \overline{1,s}$; $\|t_n\| = \sqrt{t_{1n}^2 + \ldots + t_{sn}^2}$ is the Euclidean norm of t_n; $\Longrightarrow N_s\{\mu, \sigma\}$ is the symbol of weak convergence of sequence of random variables to the s-dimensional normal random variable with mean $\mu = (\mu_1, \ldots, \mu_s)$ and symmetric covariance matrix $\sigma = ||\sigma_{ij}||$, $0 < \sigma_{jj} = \sigma_{jj}(x) < \infty$, $j = \overline{1,s}$; $\Re$ is the set of integers.

D e f i n i t i o n. The function $H(t) : R^s \to R^1$ and the sequence $\{H(t_n)\}$ are said to belong to the class $\mathcal{N}_{\nu,s}(t;\gamma)$, provided that

1) there exists an ε-neighborhood $\{z : |z_i - t_i| < \varepsilon;\ i = \overline{1,s}\}$, in which the function $H(z)$ and all its partial derivatives $\dfrac{\partial H(z)}{\partial z_j}$ up to the order ν are continuous and bounded;

2) for any values of variables $X_1, ..., X_n$ the sequence $\{H(t_n)\}$ is dominated by a numerical sequence $C_0 d_n^{\gamma}$, such that $d_n \uparrow \infty$, as $n \to \infty$, and $0 \le \gamma < \infty$.

Theorem 1 [6]. Let the conditions

1) $H(z), \{H(t_n)\} \in \mathcal{N}_{2,s}(t,\gamma)$,

2) $\mathbf{E}||t_n - t||^i = O\left(d_n^{-i/2}\right)$

hold for all $i \in \Re$. Then, for every $k \in \Re$

$$\left|\mathbf{E}\left[H(t_n) - H(t)\right]^k - \mathbf{E}\left[\nabla H(t)(t_n - t)^T\right]^k\right| = O\left(d_n^{-(k+1)/2}\right). \tag{10}$$

If in formula (10) $k = 1$, we obtain the principal term $\mathbf{E}\left[\nabla H(t)(t_n - t)^T\right]$ of the bias $\mathbf{E}\left[H(t_n) - H(t)\right]$ for $H(t_n)$, and at $k = 2$, we have the principal term $\mathbf{E}\left[\nabla H(t)(t_n - t)^T\right]^2$ of the mean squared error (MSE) $\mathbf{E}\left[H(t_n) - H(t)\right]^2$.

Theorem 2 (The usual central limit theorem) [1]. If $\xi_1, ..., \xi_n, ...$ is a sequence of independent and identically distributed s-dimensional vectors,

$$\mathbf{E}\xi_k = 0, \quad \sigma(x) = \mathbf{E}\{\xi_k^T \xi_k\}, \quad t_n = \frac{1}{n}\sum_{k=1}^{n} \xi_k,$$

then, as $n \to \infty$,

$$\sqrt{n} t_n \Longrightarrow N_s\{0, \sigma(x)\}.$$

Theorem 3 [6]. If $q_n(t_n - t) \Longrightarrow N_s\{\mu, \sigma\}$ for some number sequence $q_n \uparrow \infty$, the function $H(z)$ is differentiable at the point μ, $\nabla H(\mu) \neq 0$, then

$$q_n\left(H(t_n) - H(\mu)\right) \Longrightarrow N_1\{\nabla H(\mu)\,\mu^T,\ \nabla H(\mu)\,\sigma\,\nabla H^T(\mu)\}.$$

6 Bias and MSE of Estimator $\hat{\overline{A}}_{x_1:\ldots:x_m}$

Here, we will obtain the principal term of the asymptotic MSE and the bias convergence rate of estimator (3).

Theorem 4. If the survival function $S(x_1, \ldots, x_m) > 0$ and $S(t_1, \ldots, t_m)$ is continuous at a point $(x_1, \ldots, x_m)$, then

1) for the bias $b\left(\hat{\overline{A}}_{x_1:\ldots:x_m}\right)$ of estimator (3) we have

$$\left|b\left(\hat{\overline{A}}_{x_1:\ldots:x_m}\right)\right| = \left|\mathbf{E}\hat{\overline{A}}_{x_1:\ldots:x_m} - \overline{A}_{x_1:\ldots:x_m}\right| = O\left(n^{-1}\right);$$

2) the MSE $u^2\left(\hat{\overline{A}}_{x_1:\ldots:x_m}\right)$ is given by the formula

$$u^2(\hat{\overline{A}}_{x_1:\ldots:x_m}) = \mathbf{E}\left(\hat{\overline{A}}_{x_1:\ldots:x_m} - \overline{A}_{x_1:\ldots:x_m}\right)^2 =$$

$$= \frac{\Phi(x_1,\ldots,x_m,2\delta)\,S(x_1,\ldots,x_m) - \Phi^2(x_1,\ldots,x_m,\delta)}{n\,S^3(x_1,\ldots,x_m)} + O\left(\frac{1}{n^{3/2}}\right).$$

Proof. For estimator $\hat{\overline{A}}_{x_1:\ldots:x_m}$ (3) in the notation of Theorem 1 we have:

$$s=2,\quad t_n=(t_{1n},t_{2n})=(\Phi_n(x_1,\ldots,x_m,\delta),S_n(x_1,\ldots,x_m)),$$

$$d_n=n,\quad H(t_n)=\frac{\Phi_n(x_1,\ldots,x_m,\delta)}{S_n(x_1,\ldots,x_m)}=\hat{\overline{A}}_{x_1:\ldots:x_m},$$

$$t=(t_1,t_2)=(\Phi(x_1,\ldots,x_m,\delta),S(x_1,\ldots,x_m)),\quad H(t)=\frac{t_1}{t_2}=\overline{A}_{x_1:\ldots:x_m},$$

$$H_1(t)=\frac{1}{S(x_1,\ldots,x_m)},\quad H_2(t)=-\frac{\Phi(x_1,\ldots,x_m,\delta)}{S^2(x_1,\ldots,x_m)},\quad \nabla H(t)=(H_1(t),H_2(t))\neq 0.$$

.

The sequence $\{H(t_n)\}$ satisfies the condition 1) of Theorem 1 with $C_0=1$ and $\gamma=0$. Indeed, according to (3)

$$H(t_n)=\frac{\Phi_n(x_1,\ldots,x_m,\delta)}{S_n(x_1,\ldots,x_m)}=\frac{\frac{1}{n}\sum\limits_{i=1}^{n}e^{-\delta\prod_{j=1}^{m}Z_{ji}}\prod\limits_{j=1}^{m}I(Z_{ji}>0)}{\frac{1}{n}\sum\limits_{i=1}^{n}\prod\limits_{j=1}^{m}I\left(Z_{ji}>0\right)}\le 1. \tag{11}$$

Further, in view of $t_2=S(x_1,\ldots,x_m)>0$ the function $H(t)$ satisfies the condition 1) of Theorem 1 . Also, this function satisfies the condition 2) of Theorem 1 due to Lemma 3.1 [5], as for all $i\in\Re$ such inequalities hold:

$$\mathbf{E}\prod_{j=1}^{m}I^i\left(Z_j>0\right)=S(x_1,\ldots,x_m)\le 1,$$

$$\mathbf{E}e^{-i\delta\prod_{j=1}^{m}Z_j}\prod_{j=1}^{m}I^i(Z_j>0)\le S(x_1,\ldots,x_m)\le 1.$$

Therefore,

$$\mathbf{E}|\Phi_n(x_1,\ldots,x_m,\delta)-\Phi(x_1,\ldots,x_m,\delta)|^i=O\left(n^{-\frac{i}{2}}\right),$$

$$\mathbf{E}|S_n(x_1,\ldots,x_m) - S_n(x_1,\ldots,x_m)|^i = O\left(n^{-\frac{i}{2}}\right).$$

Taking $k=1$ in formula (10), we get

$$\mathbf{E}\{\hat{\overline{A}}_{x_1:\ldots:x_m}\} = \overline{A}_{x_1:\ldots:x_m} + \frac{1}{t_2}\,\mathbf{E}\{t_{1n}-t_1\} - \frac{t_1}{t_2^2}\,\mathbf{E}\{t_{2n}-t_2\} + O\left(\frac{1}{n}\right).$$

Since functions $t_1 = t_1(x_1,\ldots,x_m)$, $t_2 = t_2(x_1,\ldots,x_m)$ are continuous, we have $\mathbf{E}\{t_{1n}\} = t_1$, $\mathbf{E}\{t_{2n}\} = t_2$, and $\mathbf{E}\{\hat{\overline{A}}_{x_1:\ldots:x_m}\} = \overline{A}_{x_1:\ldots:x_m} + O\left(\frac{1}{n}\right)$, i.e., $\hat{\overline{A}}_{x_1:\ldots:x_m}$ is the asymptotically unbiased estimator.

Now, putting $k=2$ in (10) and taking into account unbiasedness of t_{1n}, t_{2n}, we find the formulas for the variances and covariance:

$$u^2(\hat{\overline{A}}_{x_1:\ldots:x_m}) = \frac{1}{t_2^2}\,\mathbf{D}\{t_{1n}\} + \frac{t_1^2}{t_2^4}\,\mathbf{D}\{t_{2n}\} - 2\,\frac{t_1}{t_2^3}\,\mathrm{cov}\{t_{1n},t_{2n}\} + O\left(\frac{1}{n^{3/2}}\right). \qquad (12)$$

Denote

$$\phi_i(x_1,\ldots,x_m,\delta) = e^{-\delta\,\Pi_{j=1}^m Z_{ji}} \prod_{j=1}^m I(Z_{ji}>0),$$

$s_i(x_1,\ldots,x_m) = \prod\limits_{j=1}^m I\left(Z_{ji}>0\right)$. In view of the randomness of the sample $(Z_{11},\ldots,Z_{m1})$, $\ldots$, $(Z_{1n},\ldots,Z_{mn})$, we have

$$\mathbf{D}\{t_{1n}\} = \mathbf{D}\left\{\frac{1}{n}\sum_{i=1}^n \phi_i(x_1,\ldots,x_m,\delta)\right\} = \frac{1}{n}\,\mathbf{D}\left\{\phi_1(x_1,\ldots,x_m,\delta)\right\} =$$

$$= \frac{1}{n}\,\left(\mathbf{E}\left\{\phi_1(x_1,\ldots,x_m,2\delta)\right\} - \mathbf{E}^2\left\{\phi_1(x_1,\ldots,x_m,\delta)\right\}\right) =$$

$$= \frac{1}{n}\,\left(\Phi(x_1,\ldots,x_m,2\delta) - \Phi^2(x_1,\ldots,x_m,\delta)\right),$$

$$\mathbf{D}\{t_{2n}\} = \mathbf{D}\left\{\frac{1}{n}\sum_{i=1}^n s_i(x_1,\ldots,x_m)\right\} = \frac{1}{n}\,S(x_1,\ldots,x_m)\left(1-S(x_1,\ldots,x_m)\right),$$

$$\mathrm{cov}\{t_{1n},t_{2n}\} = \frac{1}{n}\mathrm{cov}\left\{\phi_1(x_1,\ldots,x_m,\delta), s_1(x_1,\ldots,x_m)\right\} =$$

$$= \frac{1}{n}\Phi(x_1,\ldots,x_m,\delta)\left(1-S(x_1,x_2,\ldots,x_m)\right).$$

Now, we substitute the found expressions in (11) and the second assertion of the Theorem 4 has been proved.

7 Asymptotic Normality of Estimator $\hat{\overline{A}}_{x_1:\ldots:x_m}$

Theorem 5. Under the conditions of Theorem 4

$$\sqrt{n}[\hat{\overline{A}}_{x_1:\ldots:x_m} - \overline{A}_{x_1:\ldots:x_m}] \Longrightarrow$$

$$\Longrightarrow \mathsf{N}_1\left\{0, \frac{\Phi(x_1,\ldots,x_m,2\delta)\,S(x_1,\ldots,x_m) - \Phi^2(x_1,\ldots,x_m,\delta)}{S^3(x_1,\ldots,x_m)}\right\}.$$

Proof. In the notation of Theorem 2 and Theorem 3, we have: $q_n = \sqrt{n}$. Since $S(x_1,\ldots,x_m) > 0$, function $H(t) \in \mathcal{N}_{1,2}(t)$. Taking into account unbiasedness of $S_n(x_1,\ldots,x_m)$ and $\Phi_n(x_1,\ldots,x_m,\delta)$, we have $\mu^T = 0$.

According to Section 6, the elements of the covariance matrix σ are defined by the formulas:

$$\sigma_{11} = \Phi(x_1,\ldots,x_m,2\delta) - \Phi^2(x_1,\ldots,x_m,\delta),$$

$$\sigma_{12} = \sigma_{21} = \Phi(x_1,\ldots,x_m,\delta)(1 - S(x_1,\ldots,x_m)),$$

$\sigma_{22} = S(x_1,\ldots,x_m)(1 - S(x_1,\ldots,x_m))$. That is why $\nabla H(t)\,\mu^T = 0$,

$$\nabla H(t)\,\sigma\,\nabla H(t)^T = \frac{\Phi(x_1,\ldots,x_m,2\delta)\,S(x_1,\ldots,x_m) - \Phi^2(x_1,\ldots,x_m,\delta)}{S^3(x_1,\ldots,x_m)}.$$

Theorem 5 is proved.

8 Synthesis of Nonparametric Estimators of the Net Premiums in Collective Life Insurance for Other Forms of Insurance

The above considered estimators of the net premiums were constructed for whole insurance; now we will consider other forms of insurances.

- **The p-years term life insurance**

  In this case the benefit to pay if the insured will die during of the contract validity. The company does not pay the benefit if the insured has lived p years. Then

$$\hat{\overline{A}}_{\frac{[k]}{x_1:\ldots:x_m}:p]} = \frac{\frac{1}{n}\sum_{i=1}^{n} e^{-\delta Z_{(m-k+1)i}} I(0 < Z_{(m-k+1)i} \le p)}{\frac{1}{n}\sum_{i=1}^{n} I(Z_{(m-k+1)i} > 0)}.$$

- **The p-years endowment life insurance**

  Such form of insurance provides for a payment either following the death of the insured or upon his survival to the end of the p-years term. The given form of

insurance accumulates the client's capital. Then, the nonparametric estimator of the net premium is

$$\overset{\wedge}{\overline{A}}{}^{s}{}_{\underset{x_1:\ldots:x_m}{[k]}:p\rceil} = \frac{S_n(x_1,\ldots,x_m) - S_n(x_1+p,\ldots,x_m+p)}{S_n(x_1,\ldots,x_m)} \times$$

$$\times \overset{\wedge}{\overline{A}}{}_{\underset{x_1:\ldots:x_m}{[k]}:p\rceil} + \frac{S_n(x_1+p,\ldots,x_m+p)}{\overline{S}_n(x_1,\ldots,x_m)}\, e^{-\delta\, p}.$$

- **The r-years deferred life insurance**

  This form of insurance provides for a benefit following the death of the insured when he dies at least r years following policy issue. Here the net premium is expressed in the form

$$r|\overset{\wedge}{\overline{A}}{}_{\underset{x_1:\ldots:x_m}{[k]}} = \frac{\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} e^{-\delta\, Z_{(m-k+1)i}} I(r < Z_{(m-k+1)i})}{\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} I(Z_{(m-k+1)i} > 0)}.$$

9 Estimation of Joint-Life Annuity

As in the case of individual insurance [4, 30], we determine the joint-life annuity by making use of the corresponding net premium (see formulas (2)–(4), (6)):

$$\overline{a}_{x_1:\ldots:x_m}(\delta) = \frac{1}{\delta}\left(1 - \overline{A}_{x_1:\ldots:x_m}\right) = \frac{1}{\delta}\left(1 - \frac{\Phi(x_1:\ldots:x_m,\delta)}{S(x_1,\ldots,x_m)}\right). \tag{13}$$

So, in accordance with (3), we obtain the following estimator of the joint-life annuity:

$$\hat{\overline{a}}_{x_1:\ldots:x_m}(\delta) = \frac{1}{\delta}\left(1 - \frac{\Phi_n(x_1:\ldots:x_m,\delta)}{S_n(x_1,\ldots,x_m)}\right) =$$

$$= \frac{1}{\delta}\left(1 - \frac{\displaystyle\sum_{i=1}^{n} e^{-\delta\, \Pi_{j=1}^{m} Z_{ji}} \prod_{j=1}^{m} I(Z_{ji} > 0)}{\displaystyle\sum_{i=1}^{n}\prod_{j=1}^{m} I\left(Z_{ji} > 0\right)}\right). \tag{14}$$

Find the principal term of the asymptotic MSE and the bias convergence rate of estimator (13).

Theorem 6. If the survival function $S(x_1,\ldots,x_m) > 0$ and $S(t_1,\ldots,t_m)$ is continuous at a point $(x_1,\ldots,x_m)$, then

1) for the bias $b\left(\hat{\overline{a}}_{x_1:\ldots:x_m}(\delta)\right)$ of estimator (13) we have

$$\left|b\left(\hat{\overline{a}}_{x_1:\ldots:x_m}(\delta)\right)\right| = \left|\mathbf{E}\hat{\overline{a}}_{x_1:\ldots:x_m}(\delta) - \overline{a}_{x_1:\ldots:x_m}(\delta)\right| = O\left(n^{-1}\right);$$

2) the MSE $u^2\left(\hat{\overline{a}}_{x_1:\ldots:x_m}(\delta)\right)$ is given by the formula

$$u^2(\hat{\overline{a}}_{x_1:\ldots:x_m}(\delta)) = \mathbf{E}\left(\hat{\overline{a}}_{x_1:\ldots:x_m}(\delta) - \overline{a}_{x_1:\ldots:x_m}(\delta)\right)^2 =$$

$$= \frac{\Phi(x_1,\ldots,x_m,2\delta)\,S(x_1,\ldots,x_m) - \Phi^2(x_1,\ldots,x_m,\delta)}{n\,\delta^2\,S^3(x_1,\ldots,x_m)} + O\left(\frac{1}{n^{3/2}}\right).$$

Proof. For estimator $\hat{\overline{a}}_{x_1:\ldots:x_m}(\delta)$ (13) in the notation of Theorem 1 we have:

$$s=2,\quad d_n=n,\quad t_n=(t_{1n},t_{2n})=(\Phi_n(x_1,\ldots,x_m,\delta),\,S_n(x_1,\ldots,x_m)),$$

$$H(t_n)=\frac{1}{\delta}\left(1-\frac{\Phi_n(x_1,\ldots,x_m,\delta)}{S_n(x_1,\ldots,x_m)}\right)=\hat{\overline{a}}_{x_1:\ldots:x_m}(\delta),$$

$$t=(t_1,t_2)=(\Phi(x_1,\ldots,x_m,\delta),\,S(x_1,\ldots,x_m)),\ H(t)=\frac{1}{\delta}\left(1-\frac{t_1}{t_2}\right)=\overline{a}_{x_1:\ldots:x_m}(\delta),$$

$$H_1(t)=\frac{1}{\delta S(x_1,\ldots,x_m)},\ \ H_2(t)=-\frac{\Phi(x_1,\ldots,x_m,\delta)}{\delta S^2(x_1,\ldots,x_m)},\ \ \nabla H(t)=(H_1(t),H_2(t))\neq 0.$$

The sequence $\{H(t_n)\}$ satisfies the condition 1) of Theorem 1 with $C_0=\dfrac{1}{\delta}$ and $\gamma=0$. Taking into account (13) and the inequalities $0\le\dfrac{\Phi_n(x_1,\ldots,x_m,\delta)}{S_n(x_1,\ldots,x_m)}\le 1$ (see (11)), we have

$$H(t_n)=\frac{1}{\delta}\left(1-\frac{\Phi_n(x_1,\ldots,x_m,\delta)}{S_n(x_1,\ldots,x_m)}\right)\le\frac{1}{\delta}.$$

Further, the proof is carried out similarly to the proof of Theorem 4 and therefore is not given.

Theorem 7. Under the conditions of Theorem 4

$$\sqrt{n}[\hat{\overline{a}}_{x_1:\ldots:x_m}(\delta)-\overline{a}_{x_1:\ldots:x_m}(\delta)]\Longrightarrow$$

$$\Longrightarrow \mathsf{N}_1\left\{0,\frac{\Phi(x_1,\ldots,x_m,2\delta)\,S(x_1,\ldots,x_m)-\Phi^2(x_1,\ldots,x_m,\delta)}{\delta^2 S^3(x_1,\ldots,x_m)}\right\}.$$

Conclusions

The paper deals with the estimation problem of the current values of net premiums and life annuities. The asymptotic properties of the estimators are proved: unbiasedness, consistency and normality. The principal terms of the asymptotic MSEs of the proposed estimators are found. Statistical modeling within the framework of the de Moivre model shows that the quality of estimation according to empirical criterion improves with the growth of the sample size. Note that the improved estimators of net premiums and life annuities can be obtained by substituting of empirical survival functions by the smooth empirical survival functions (cf. [2, 3, 9, 11, 15, 17, 24], [27]–[29], [8, 36, 38, 40, 41, 43, 44]).

References

[1] Aase K.K. (1993). Premiums in a Dynamic Model of a Reinsurance Market. *Scandinavian Actuarial Journal*. No. **2**, pp. 134-160.

[2] Altman N., Leger C. (1995) Bandwidth Selection for Kernel Distribution Function Estimation. *J. Statist. Plann. Inference*. Vol. **46**, pp. 195-214.

[3] Azzalini A. (1981) A Note on the Estimation of a Distribution Function and Quantiles by a Kernel Method. *Biometrika*. Vol. **68**. No. **1**, pp. 326-328.

[4] Arora N., Arora P. (2014). Insurance Premium Optimization: Perspective of Insurance Seeker and Insurance Provider. *Journal of Management and Science*. Vol. **4**. No. **1**, pp. 43-53.

[5] Bloink R. (2010). Premium Financed Surprises: Cancellation of Indebtedness Income and Financed Life Insurance. *Tax Lawyer*. Vol. **63**. No. **3**, pp. 223-227.

[6] Bohnert A., Gatzert N. (2012). Analyzing Surplus Appropriation Schemes in Participating Life Insurance from the Insurer's and the Policyholder's Perspective. *Insurance: Mathematics and Economics*. Vol. **50**. No. **1**, pp. 64-78.

[7] Borovkov A.A. (1998) *Mathematical Statistics*. Gordon & Breach, New York.

[8] Bowers N.L., Gerber H.U., Hickman J.C., Jones D.A., Nesbitt C.J. (1986). *Actuarial Mathematics*. The Society of Actuaries, Itasca, Illinois.

[9] Bowman A., Hall P., Prvan T. (1998) Trust Bandwidth Selection for the Smoothing of Distribution Functions. *Biometrika*. Vol. **85**. No. **4**, pp. 799-808.

[10] Buhlmann H. (1984). The General Economic Premium Principle. *ASTIN Bulletin*. Vol. **14**, pp. 13-21.

[11] Chu I.S. (1995). Bootstrap Smoothing Parameter Selection for Distribution Function Estimation. *Math. Japon*. Vol. **41**. No. **1**, pp. 189-197.

[12] Dobrovidov A.V., Koshkin G.M. (1997). *Nonparametric Signal Estimation*. Nauka. Fizmatlit, Moskva (in Russian).

[13] Dong Y. (2011). Fair Valuation of Life Insurance Contracts under a Correlated Jump Diffusion Model. *ASTIN Bulletin*. Vol. **41**. No. **2**, pp. 429-447.

[14] Falin G.I. (2002). *Mathematical Foundations of the Theory of Life Insurance and Pension Schemes*. Ankil Publ., Moscow (in Russian).

[15] Falk M. (1983) Relative Efficiency and Deficiency of Kernel Type Estimators of Smooth Distribution Functions. *Statist. Neerlandica*. Vol. **37**, pp. 73-83.

[16] Faust R., Schmeiser H., Zemp A. (2013). A Performance Analysis of Participating Life Insurance Contracts. *Insurance: Mathematics and Economics*. Vol. **51**. No. **1**, pp. 158-171.

[17] Fuks I., Koshkin G. (2015). Smooth Recurrent Estimation of Multivariate Reliability Function. *Proceedings. The International Conference on Information and Digital Technologies (IDT 2015)*. Zilina, Slovakia, pp. 84-89.

[18] Gatzert N. (2008). Asset Management and Surplus Distribution Strategies in Life Insurance: An Examination with Respect to Risk Pricing and Risk Measurement. *Insurance: Mathematics and Economics*. Vol. **42**. No. **2**, pp. 839-849.

[19] Gatzert N., Wesker H. (2012). The Impact of Natural Hedging on a Life Insurer's Risk Situation. *Journal of Risk Finance*. Vol. **13**. No. **5**, pp. 396-423.

[20] Gerber H.U. (1980). Principles of Premium Calculation and Reinsurance. *Transactions of the 21st International Congress of Actuaries*, pp. 137-142.

[21] Gerber H. (1997). *Life Insurance Mathematics*. 3rd ed. Springer-Verlag, New York.

[22] Huang H., Lee Y.-T. (2010). Optimal Asset Allocation for a General Portfolio of Life Insurance Policies. *Insurance: Mathematics and Economics*. Vol. **46**. No. **2**, pp. 271-280.

[23] Ibragimov I.A., Khasminskii R.Z. (1981). *Statistical Estimation: Asymptotic Theory*. Springer, Berlin; New York.

[24] Jones M.C. (1990) The Performance of Kernel Density Functions in Kernel Distribution Function Estimation. *Statist. Probab. Lett.* Vol. **9**, pp. 129-132.

[25] Koshkin G.M. (1998). Nonparametric Estimation of Moments of Net Premium in Life Insurance. *Computer Data Analysis and Modeling: Proceedings of the Fifth International Conference*. Vol. **1**. BSU, Minsk, pp. 148-153.

[26] Koshkin G.M. (1999). Deviation Moments of the Substitution Estimator and Its Piecewise Smooth Approximations. *Siberian Mathematical Journal*. Vol. **40**. No. **3**, pp. 515-527.

[27] Koshkin G.M. (2014) Smooth Estimators of the Reliability Functions for Non-Restorable Elements. *Russian Physics Journal*. Vol. **57**. No. **5**, pp. 672-681.

[28] Koshkin G.M. (2015) Smooth Recurrent Estimators of the Reliability Functions. *Russian Physics Journal*. Vol. **58**. No. **7**, pp. 1018-1025.

[29] Koshkin G.M. (2016) Smooth Nonparametric Estimation of the Failure Rate Function and Its First Two Derivatives. *Russian Physics Journal*. Vol. **59**. No. **6**, pp. 833-844.

[30] Koshkin G.M. Gubina O.V. (2016). Estimation of the present values of life annuities for the different actuarial models. *Proceedings. The Second International Symposium on Stochastic Models, in Reliability Engineering, Life Science, and Operations Management*. The Institute of Electrical and Electronics Engineers, Beer Sheva, Israel, pp. 506-510.

[31] Koshkin G.M., Lopukhin Ya.N. (2001). On Estimation of Net Premium in Collective Life Insurance. *The 5th Korea-Russian International Symposium on Science and Technology. Proceedings. KORUS 2001*. Vol. **2**. Tomsk Polytechnic University, Tomsk, pp. 296-299.

[32] Koshkin G.M., Lopukhin Ya.N. (2001). Estimation of Net Premiums in Collective Models of Life Insurance. *XIth Annual International AFIR Colloquium*. Vol. **2**. Canadian Institute of Actuaries, Toronto, pp. 447-457.

[33] Koshkin G.M., Lopukhin Ya.N. (2001). Nonparametric Estimation of Net Premiums in Collective Insurance. *Computer Data Analysis and Modeling: Robustness and Computer Intensive Methods: Proceedings of the Sixth International Conference*. Vol. **1**. BSU, Minsk, pp. 236-241.

[34] Maurer R., Rogalla R., Siegelin I. (2013). Participating Payout Life Annuities: Lessons from Germany. *ASTIN Bulletin*. Vol. **43**. No. **2**, pp. 159-187.

[35] Nadaraya E.A. (1964). Some New Estimates of Distribution Function. *Theory of Probability and its Applications*. Vol. **9**. No. **3**, pp. 497-500.

[36] Reiss R.-D. (1981). Nonparametric Estimation of Smooth Distribution Functions. *Scand. J. Statist.* Vol. **8**, pp. 116-119.

[37] Rocio P., Aguilar C., Xu C. (2010). Design Life Insurance Participating Policies Using Optimization Techniques, International Journal of Innovative Computing. *Information and Control*. Vol. **6**. No. **4**, pp. 1655-1666.

[38] Sarda P. (1993) Smoothing Parameter Selection for Smooth Distribution Functions. *J. Statist. Plann. Inference Inf.* Vol. **35**, pp. 65-75.

[39] Schmeiser H., Wagner J. (2011). A Joint Valuation of Premium Payment and Surrender Options in Participating Life Insurance Contracts. *Insurance: Mathematics and Economics*. Vol. **49**. No. **3**, pp. 580-596.

[40] Shao Y. Xiang X. (1997). Some Extensions of the Asymptotics of a Kernel Estimator of a Distribution Function. *Statist. Probab. Lett.* Vol. **34**, pp. 301-308.

[41] Shirahata S. Chu I.S. (1992). Integrated Squared Error of Kernel-Type Estimator of Distribution Function. *Ann. Inst. Statist. Math.* Vol. **44**. No. **3**, pp. 579-591.

[42] Sliwinski A., Michalski T., Roszkiewicz M. (2013). Demand for Life Insurance - An Empirical Analysis in the Case of Poland. *The Geneva Papers on Risk and Insurance*. Vol. **38**, pp. 62-87.

[43] Swanepoel J.W.H. (1988). Mean Integrated Squared Error Properties and Optimal Kernels When Estimating a Distribution Function. *Comm. Statist. Theory Methods*. Vol. **17**. No. **11**, pp. 3785-3799.

[44] Una-Alvarez J., Gonzalez-Manteiga W., Cadarso-Suarez C. (2000). Kernel Distribution Function Estimation under the Koziol-Green Model. *J. Statist. Plann. Inference*. Vol. **87**, pp. 199-219.

Method of the estimation of uncertainties in multiparameter measurements of correlated quantities

Zygmunt Lech Warsza[1] and Jacek Puchalski[2]
[1] *Industrial Research Institute of Automation and Measurements (PIAP), Warszawa Poland*
[2] *Central Office of Measures (GUM), Warszawa Poland*
e-mail: zlw1936@gmail.com, jacek.puchalski@gum.gov.pl

Abstract

The paper presents an improved version of the method of evaluation the multiparameter measurement uncertainties stated in the Supplement 2 to GUM guide. This was done on the example of two-parameter jointed measurements. It consists the correlation of individual components of the type A and/or type B uncertainty of measurands. The general formulas for the covariance matrix, final uncertainties and correlation coefficient were determined as well as formulas for several specific cases, presented in Table 1. The graphs show the correlation coefficients of the output quantities as a function of the type B contribution in the uncertainty of the input quantities. Also, there are provided examples of estimation of the uncertainty and correlation coefficients for the sum and difference on the example of two temperatures. It has been demonstrated that the inclusion of correlations of uncertainty components makes the uncertainty evaluations more reliable and accurate.

Keywords: multivariate measurements, correlations of the type A and type B uncertainty components, vector propagation of variance, resultant correlation coefficient.

Introduction

In the GUM guide, a concept called "measurement uncertainty" (MU) was introduced to estimate the accuracy of measurements. It is the width of the interval, and for multi-parameter measurements - a description of the boundaries of the so-called coverage area [3], [11], in which the estimator of the value of measurand, obtained after processing of the measurement results, can occur with a certain probability. The measurement uncertainty assessment is based on the determination of its type A and type B components, designated as u_A and u_B, respectively [1]. They are defined as standard deviations of the resultant distribution with a probability density function $p(x)$ constituting a convolution of two statistically independent density distributions $p(x_A)$ and $p(x_B)$. The first one describes a random spread of the values of experimental observations obtained experimentally. The second one is a hypothetical distribution randomizing the assumed changes in the value of many systematic errors

of different origin, unknown in value during the measurement experiment. These errors may occur randomly in the long-term use of the measuring object, of instrument or measuring system and in various permissible ambient conditions other than during provided control measurements. The standard deviation of the $p(x)$ distribution, i.e. its uncertainty u, is the geometric sum of component uncertainties u_A and u_B, i.e.:

$$u = \sqrt{u_A^2 + u_B^2}. \tag{1}$$

The uncertainty of any single quantity described by (1) has the same form as for the uncertainty of the sum of two uncorrelated variables, when values of one of them have only type A, and the other - type B uncertainty.

If a function describing the density distribution of $p(x)$ is known, e.g. the Gaussian function for a normal distribution, the expanded uncertainty U with the required coverage probability, e.g. 95 % or 99 %, is determined analytically basing on standard deviation u and the expansion coefficient. For other randomized distributions in experiments, the expanded uncertainty U is also determined by the numerical Monte Carlo method according to Supplement 1 [2] to the GUM guide.

Standard uncertainty u_A of component type A depends on the distribution of the measurement observations and it is determined statistically. On the other hand, due to u_B uncertainty, the possible impacts of many values affecting measurement results of unknown values, irremovable from the sample observation, are randomized, as there are no data to calculate the corrections for them. Therefore, u_B values can only be estimated heuristically based on knowledge about predicted ranges and distributions of values of influencing quantities and their interaction functions. In the known operating conditions with a limited range of changes, the uncertainty component u_B may be significantly smaller than that assumed during the calibration of the device for the full permissible conditions of its application [8], [9].

Measurement uncertainty (MU) is applicated as a basic element of conformity assessment of products with requirements [4-6]. It allows to compare the results of various tests, check the exceeding of limits or meet the tolerance requirements of products. Thanks to the possibility of comparing the calibration results with the requirements, it is also the basis for metrological acceptance of the measuring equipment. Laboratories accredited according to ISO / IEC 17025 [6] must calculate the MU for each test method used in the granted accreditation scope. There is, however, a certain discrepancy between the description of measurement accuracy through uncertainty with a certain probability and description of the accuracy of measuring devices and devices by the permissible maximum errors, i.e. border errors.

The metrological properties of the measuring equipment and the parameters of many devices and processes that fulfill the responsible functions, besides careful calibration, also require periodic inspection during the period of their operation [5, 16]. For this purpose, measurement experiments are carried out, including the study of these parameters as single and multiparameter combined monolayers. They are rarely performed by one, and usually by many instruments, under the same or different environmental conditions and at the spread of the values of repeated measurement

observations caused by common or different causes. This practical multi-variant situation creates different relationships between individual type A and type B component uncertainties of measured quantities. Many such situations do not include recommendations for estimating uncertainty of measurements, given in GUM [1] and all its Supplements [2-4] and in other international regulations, e.g. [5], [6]. Zakharow in [7] considered the non-heuristic estimates of the correlation of the quantities influencing the uncertainty of single parameter measurements, called "logical ". In the literature on multiparameter measurements, we have not found an analysis of impact of the correlation of uncertainty components type A and/or type B of the input quantities.

The purpose of this publication is to develop a method for determining the relationship between the uncertainty of the measurement results of several output values, when the individual type A and B type uncertainties of few input quantities are correlated. This is a proposal to extend the scope of the GUM guide and its Supplement 2 applications.In this work the impact of various cases of correlations of uncertainty components of directly measured quantities on the resultant uncertainty of the quantities indirectly determined from measurement data will be analyzed. This will be illustrated by the examples of two-parameter measurements described by a linear function and several simple non-linear functions. Appropriate formulas that take into account the correlation of uncertainty components will be determined. This is particularly important when the results of multi-parameter measurements are used later together in various applications.

1 Uncertainty of 2D function of measurand with correlated uncertainty components

The mathematical model of the propagation of uncertainties given in Supplement 2 [3] to the GUM guide [1] is used to determine the uncertainty of components for multidimensional indirect measurements. In such measurements, the m-dimensional output measurand $\boldsymbol{Y}$ depends on data of measured directly n-dimensional measurand $\boldsymbol{X}$ by general equation

$$\boldsymbol{Y} = \boldsymbol{F}(\boldsymbol{X}). \tag{2}$$

where: $\boldsymbol{Y}$ and $\boldsymbol{X}$ - vectors with elements that are values of output and input variables.

Propagation of the uncertainty of these measurements is described as a matrix relation between the covariance matrices of input $\boldsymbol{X}$ and output $\boldsymbol{Y}$ variables. It occurs in the linearization of the functional $\boldsymbol{F}$ in the formula (2) through derivatives and has the form [3]:

$$\boldsymbol{U_Y} = \boldsymbol{S} \cdot \boldsymbol{U_X} \cdot \boldsymbol{S^T} \tag{3}$$

where*:* $\boldsymbol{S}$ - the sensitivity matrix);; $\boldsymbol{U_X}$, $\boldsymbol{U_Y}$ - covariance matrix of input $\boldsymbol{X}$ and

output $\boldsymbol{Y}$ vectors escribed as follows

$$\boldsymbol{S}=\begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix},\ \boldsymbol{U_X}=\begin{bmatrix} u_{x1}^2 & \cdots & \rho_{x1n}u_{x1}u_{xn} \\ \cdots & \cdots & \cdots \\ \rho_{xn1}u_{xn}u_{x1} & \cdots & u_{xn}^2 \end{bmatrix}, \qquad (3a,b)$$

$$\boldsymbol{U}_Y = \begin{bmatrix} u_{y1}^2 & \cdots & \rho_{y1m}u_{y1}u_{ym} \\ \cdots & \cdots & \cdots \\ \rho_{y1m}u_{ym}u_{y1} & \cdots & u_{ym}^2 \end{bmatrix} \qquad (3c)$$

If the $\boldsymbol{Y}$ vector will be further processed by any other function:

$$\boldsymbol{Z} = \boldsymbol{G}(\boldsymbol{Y}). \qquad (4)$$

then for the indirectly determined values of the multivariable $\boldsymbol{Z} = [z_1, \ldots, z_m]^T$ with the uncertainties $u_{z1}, \ldots, u_{zm}$, we get then extmatrix equation similar as (3) describing the propagation of uncertainties

$$\boldsymbol{U_Z} = \boldsymbol{S_G} \cdot \boldsymbol{U_Y} \cdot \boldsymbol{S_G}^{\boldsymbol{T}} \qquad (5)$$

There are many possibilities to correlate each of the two component uncertainties with their equivalents of other input quantities measured in a joint experiment. The results of both directly measured values of $\boldsymbol{X}$, as well as indirectly measured values of $\boldsymbol{Y}$ and $\boldsymbol{Z}$, are determined from the collections of repeated observations obtained in randomly variable ambient conditions. If some of these conditions are constant or variable deterministically during measurements in a known manner, only then the corrections can be made to eliminate the resulting known systematic errors. Different values of the examined quantities are measured with different meters, or even the same quantity on different ranges of one meter. There are various permissible errors of these instruments, and therefore different values of the uncertainty component type B [8-12]. In addition, the dispersion of observation values depends not only on parameters describing changes of the measurement system and the test object from ambient conditions, but also on the random and aging changes of their internal parameters. Measurement samples from the same physical object under test (e.g. in chemical analysis) in different experiments carried out under the same conditions, i.e. with the same u_B uncertainty, may have different values of uncertainty u_A.

2 Uncertainty of the different correlation cases of two quantities

In multi-parameter measurements, for measured values of elements of the input variable $\boldsymbol{X}$ different types of correlation of components of each uncertainty type A and type B may occur. This influences on the accuracy of the results of the output quantities $\boldsymbol{Y}$. Measurements of only the components of the vector $\boldsymbol{X}$ should be treated as a special case of multiparameter measurements related to each other only in a

measurement object, i.e. when $\boldsymbol{Y} = \boldsymbol{X}$. Below, some examples of determining the uncertainty of vector $\boldsymbol{Y}$ components with two values y_1, y_2 obtained from measurements of two values y_1, y_2 (or one quantity in two different experiments and under different influencing conditions) will be considered. Both input and output measurands are two-parameter (2D).For 2D measurement of input variable $\boldsymbol{X} = [x_1, x_2]$, i.e., when $\boldsymbol{Y} = \boldsymbol{X}$, i.e.: $y_1 = x_1, y_2 = x_2$, two equations (7a, b) and the output covariance matrix $\boldsymbol{U}_Y \equiv \boldsymbol{U}_{x1,2}$ will be obtained

$$\boldsymbol{U}_{x1,2} = \begin{bmatrix} u_{1A}^2 + u_{1B}^2 & \rho_A u_{1A} u_{2A} + \rho_B u_{1B} u_{2B} \\ \rho_A u_{1A} u_{2A} + \rho_B u_{1B} u_{2B} & u_{2A}^2 + u_{2B}^2 \end{bmatrix} \tag{6}$$

In the covariance matrix $\boldsymbol{U}_{x1,2}$ there are uncertainties according to the rule of totalizing variance, i.e. as sum of squares of uncertainty type A and type B in both measurements:

$$u_{x1}^2 = u_{1A}^2 + u_{1B}^2, \qquad u_{x2}^2 = u_{2A}^2 + u_{2B}^2 \tag{7a, b}$$

The correlation coefficient between variablesx_1 and x_2 is:

$$\rho_{x1,2} = \frac{\rho_A u_{1A} u_{2A} + \rho_B u_{1B} u_{2B}}{\sqrt{u_{1A}^2 + u_{1B}^2}\sqrt{u_{2A}^2 + u_{2B}^2}} \tag{8}$$

Relations between the uncertainty components of type A and of type B are illustrated in Figure 1.

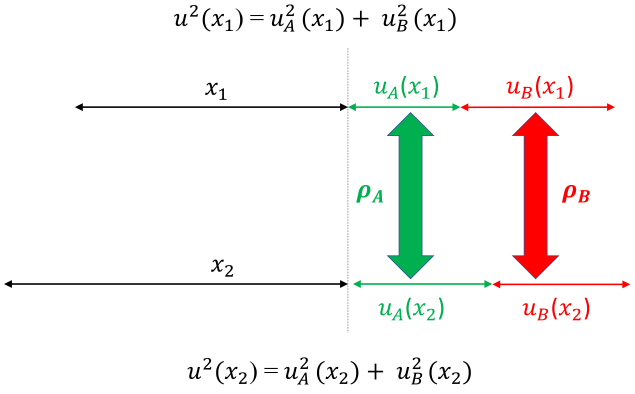


Figure 1: The relations between uncertainties of Type A and/or Type B for the 2D correlated measurand

For the uncertainty ratios of components and standard uncertainties we put given below designations: $k_{1\mathrm{B}} \equiv \frac{u_{1B}}{u_{x1}} \leq 1$, $0 \leq k_{2\mathrm{B}} \equiv \frac{u_{2\mathrm{B}}}{u_{x2}} \leq 1$. From that $\frac{u_{1A}}{u_{x1}} \equiv \sqrt{1 - k_{1B}^2}$, $\frac{u_{2\mathrm{A}}}{u_{x2}} \equiv \sqrt{1 - k_{2B}^2}$, and a simpler form of the pattern (8) is obtained

$$\rho_{x1,2} = \rho_A \sqrt{1 - k_{1B}^2}\sqrt{1 - k_{2B}^2} + \rho_B k_{1\mathrm{B}}\ k_{2\mathrm{B}} \tag{9}$$

For both quantities measured in this experiment, the coefficient $\rho_{y1,2}$ in (9) depends only on the uncertainty ratio $k_{i\mathrm{B}}$, because for each of them there is $\left(\frac{u_{iB}}{u_{xi}}\right)^2+\left(\frac{u_{iA}}{u_{xi}}\right)^2 = k_{iB}^2+(1-k_{iB}^2)=1$.

The linear condition $\frac{u_{iB}}{u_{xi}}+\frac{u_{iA}}{u_{xi}} = k_{i\mathrm{B}}+\sqrt{1-k_{iB}^2}=1$ is satisfied only for $k_{i\mathrm{B}}=0$ or $k_{i\mathrm{B}}=1$, i =1, 2.

Some specific cases of the correlation coefficient $\rho_{y1,2}$ for the results of two measurement experiments with different uncertainty of components u_{1A}, u_{2A} and u_{1B}, u_{2B} and their correlation coefficients ρ_A and ρ_B are given in Table 1. It also contains extreme cases for a combination of values $\rho_A=(0,\ 1)\,;\rho_B=(0,\ 1)$.In a special case, when $\rho_A=0$, $\rho_B=1$ and $k_{1B}^2\approx k_{1B}^2\approx 0.5$, we get: $u_{x1}=u_{x2}=\sqrt{2}u_B$ and $\rho_{y1,2}=0.5$.

Table 1. Correlation coefficient $\rho_{x1,2}$ of the measurement results of two variable measurand for different relations of its uncertainty components u_A and u_B

No	Uncertainties of the type A		Uncertainties of the type B		Correlation coefficient $\rho_{x1,2}$ between two measurands
	Values	Correlation factor	Values	Correlation factor	
	u_{1A},u_{2A}	ρ_A	u_{1B},u_{2B}	ρ_B	$\frac{\rho_A u_{1A}u_{2A}+\rho_B u_{1B}u_{2B}}{\sqrt{u_{1A}^2+u_{1B}^2}\sqrt{u_{2A}^2+u_{2B}^2}}$ $\rho_A\sqrt{1-k_{1B}^2}\sqrt{1-k_{2B}^2}$ $+\rho_B k_{1\mathrm{B}}\,k_{2\mathrm{B}}$
1.	u_{1A},u_{2A}	ρ_A	u_{1B},u_{2B}	$\rho_B=0$	$\rho_A\sqrt{1-k_{1B}^2}\sqrt{1-k_{2B}^2}$
2.	u_{1A},u_{2A}	$\rho_A=0$	u_{1B},u_{2B}	ρ_B	$\rho_B k_{1\mathrm{B}}\,k_{2\mathrm{B}}$
3.	$u_{1A}=u_{2A}=u_A$	ρ_A	$u_{1B}=u_{2B}=u_B$	ρ_B	$\rho_A(1-k_B^2)+\rho_B k_B^2$
4.	u_{1A},u_{2A}	ρ_A	$u_{1B}=u_{2B}=u_B$	$\rho_B=1$	$\rho_A\sqrt{1-k_{1B}^2}\sqrt{1-k_{2B}^2}$ $+k_{1\mathrm{B}}\,k_{2\mathrm{B}}$
5.	$u_{1A}=u_{2A}=u_A$	ρ_A=1	u_{1B},u_{2B}	ρ_B	$\sqrt{1-k_{1B}^2}\sqrt{1-k_{2B}^2}$ $+\rho_B k_{1\mathrm{B}}\,k_{2\mathrm{B}}$
6.	u_{1A},u_{2A}	$\rho_A=0$	u_{1B},u_{2B}	$\rho_B=0$	0
7.	$u_{1A}=u_{2A}=u_A$ $u_{1B}=u_{2B}=u_B$ u_A=u_B	ρ_A	$u_{1A}=u_{2A}=u_A$ $u_{1B}=u_{2B}=u_B$ u_A=u_B	ρ_B	$0.5\,(\rho_A+\rho_B)$
		ρ_A =0		ρ_B =1	$\rho_{y1,2}=\frac{1}{2}$
		ρ_A =1		ρ_B =0	
		ρ_A =1		ρ_B =1	$\rho_{y1,2}=1$

3 Graphs of the resultant correlation coefficient $\boldsymbol{\rho}_{x1,2}$

The form of the function (9) is invariant to the changes of values u_{1A} to u_{1B}, u_{2A} to u_{2B} and ρ_A to ρ_B. Then it suffices to analyze dependencies for one type of uncertainty A or B, because for the set of variables of the second type, there will be similar relationships due to the symmetry of this pattern. The general function of the formula (9) will be used to present diagrams of the correlation coefficient. Figure 2 shows 3D graphs of the resultant values of correlation coefficient $\rho_{y1,2} = f\ (k_{1B}^2,\ k_{2B}^2)$ of the two-element (2D) output measurand $\boldsymbol{Y}$=$\boldsymbol{X}$ for three pairs of correlation coefficients $\rho_A,\ \ \rho_B$ of type A or/and type B uncertainties of elements x_1, x_2 of the 2D input measurand $\boldsymbol{X}$.Graphs have the form of curvilinear planes. The cross-sections are marked on them for $k_{2B}^2 = (0.25,\ 0.5,\ 0.81)$ equal to $k_{2B} = (\ 0.5,\ \ 0.71,\ 0.9)$ are given on Fig.3a-c as 2D diagrams of coefficient $\rho_{y1,2}$ in the function of $k_{1B}^2 = \left(\frac{u_{1B}}{u_{y1}}\right)^2$ described by the formula (10)

$$\rho_{x1,2} = f(k_{1B}^2) \tag{10}$$

or as $\rho_{x1,2} = f\ (k_{1B})$, if lower non-linear scale at the bottom of each of these drawings is used.

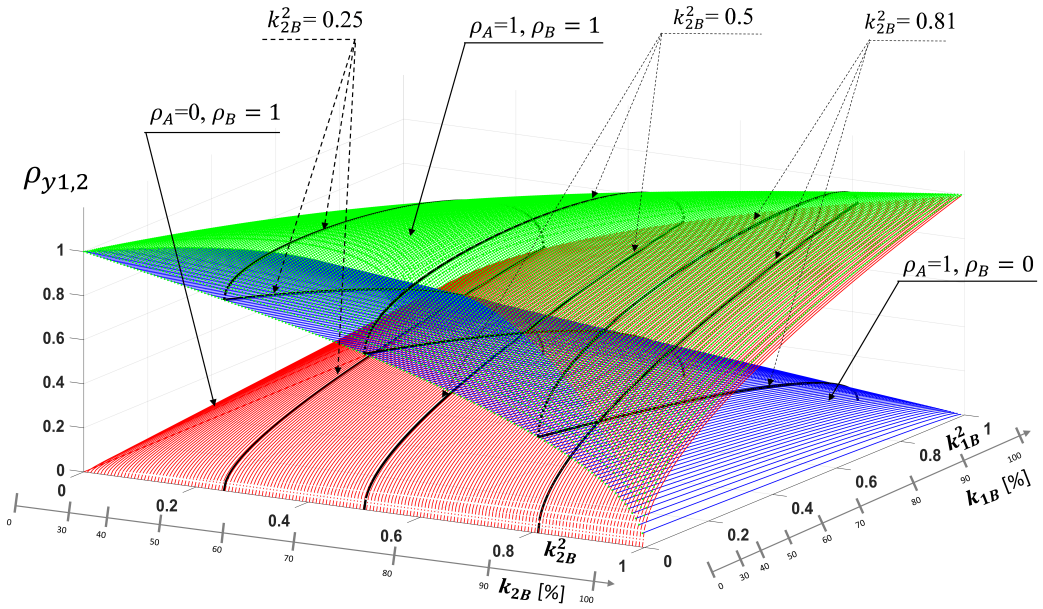


Figure 2: Relations of correlation coefficient $\rho_{y1,2} = f\ (k_{1B},\ \ k_{1B})$ of measurand $\boldsymbol{X}$ as 3D charts for three pairs of correlation coefficients of its uncertainty type A or B components*:* $\rho_A = 0, \rho_B = 1; \rho_A = 1,\ \rho_B = 1; \rho_A = 1,\ \ \rho_B = 0.$

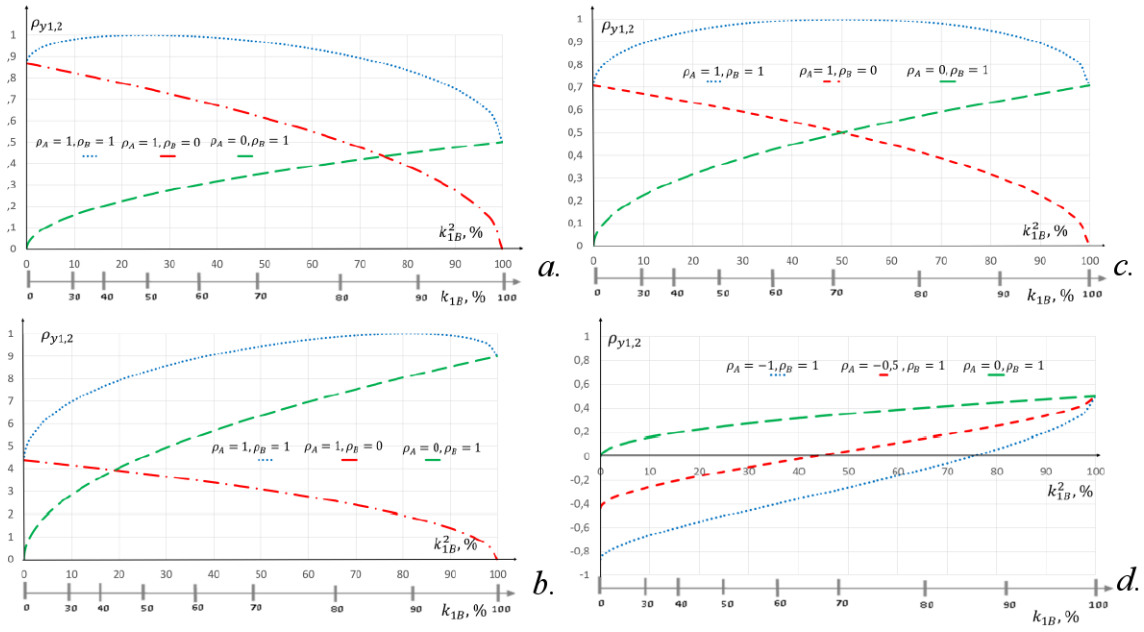


Figure 3: Correlation coefficient $\rho_{y1,2}$ between output variables as function of k_{1B}^2or $k_{1B}=\frac{u_{1B}}{u_{y1}}$ for defined values of correlation coefficients of input uncertainty components type A or/and B: $\rho_A = 0$ or 1 and $\rho_B = 0$ or 1, and values of parameters: $k_{2B}=\frac{u_{2B}}{u_{y2}}$ and k_{2B}^2 given above on **a, b, c, d** [17].

The detailed conclusions resulting from the analysis of the diagrams in Figure 3 a-d are as follows:

1. the largest correlation coefficient in the entire range k_{1B} has the curve for ρ_A=1,$\rho_B = 1$, when $k_{1B} = k_{2B}$;

2. for the value $k_{1B}^2 < 1 - k_{2B}^2$ the curve for ρ_A=1, $\rho_B = 0$ dominates over the curve ρ_A= 0, $\rho_B = 1$ and approaches the curve ρ_A= 1, $\rho_B = 1$ for the smallest values of k_{2B}^2;

3. for larger values of $k_{1B}^2 > 1 - k_{2B}^2$ the curve ρ_A=1, $\rho_B = 0$ is below the curve ρ_A=0, $\rho_B = 1$;

4. at point $k_{1B}^2 = 1 - k_{2B}^2$ curves for ρ_A=1, $\rho_B = 0$ and ρ_A=0, $\rho_B = 1$ intersect. The correlation coefficient is $\rho_{y1,2} = k_{1B}\sqrt{1 - k_{1B}^2}$ and reaches the maximum value $\rho_{y1,2} = 0.5$ for $k_{1B}^2 = 0.5$;

5. for the value of the correlation coefficient $\rho_A < 0$ (Figure 3d [17]) there a negative correlation coefficient $\rho_{y1,2}$,for the fragment of variation range of k_{1B} is obtained.For example, for ρ_A= -1, $\rho_B = 1$ and $k_{1B}^2 < 1 - k_{2B}^2$ a negative

correlation coefficient is obtained, and for $k_{1B}^2 > 1 - k_{2B}^2$ the correlation coefficient is positive. The largest range of variation k_{1B} with a negative correlation coefficient occurs for small values of k_{2B}.

On this basis, some conclusions about the uncertainty of measurements can be formulated.The correlation between the values of results of two different measurement experiments associated with each other, e.g. measurements of the same value of measurand with two different meters or on two different ranges of the same meter of different uB values, depends on theratio of the uncertainties type B and if the higher the measurement uncertainty is in relation to type A in one or both measured experiments.

The maximum correlation coefficient $\rho_{y1,2} = 1$ is achieved for the quantities of fully correlated components of type A and type B, ρ_A=1, $\rho_B = 1$, i.e. when $k_{1B} = k_{2B}$. This leads to the condition $\frac{u_{1B}}{u_{1A}} = \frac{u_{2B}}{u_{2A}}$.

The correlation coefficient increases to 1 for $k_{1B} < k_{2B}$, but decreases for $k_{1B} > k_{2B}$. For the values of $k_{1B}^2 < 1 - k_{2B}^2$ we observe a strong negative correlation for the curves ρ_A= (-1, -0.5), $\rho_B = 1$

4 Relative uncertainties

The formulas for relative uncertainties are obtained by substituting $u_i = x_i u_{ri}, u_{Ai} = x_i u_{rAi}, u_{Bi} = x_i u_{rBi}$, for i= 1,2. From the formula (8) for the correlation coefficient, we obtain:

$$\rho_{x1,2} = \frac{\rho_A u_{rA1} u_{rA2} + \rho_B u_{rB1} u_{rB2}}{\sqrt{u_{rA1}^2 + u_{rB1}^2}\sqrt{u_{rA2}^2 + u_{rB2}^2}} \tag{11}$$

Formulas of relative variances of quantities x_1, x_2 have been determined in a similar way:

$$u^2{}_r(x_1) \equiv u^2{}_{rx1} = \frac{u^2{}_{A1} + u_{B1}^2}{x_1^2} = u_{rA1}^2 + u_{rB1}^2 \tag{12a}$$

$$u^2{}_r(x_2) \equiv u^2{}_{rx2} = \frac{u^2 A2 + u_{B2}^2}{x_2^2} = u_{rA2}^2 + u_{rB2}^2 \tag{12b}$$

Relative uncertainties u_{rx1}, u_{rx2} as functions of relative uncertainties u_{rA1}, u_{rA2}, u_{rB1}, u_{rB2} are expressed as

$$u_{rx1} = \sqrt{u_{rA1}^2 + u_{rB1}^2} \quad u_{rx2} = \sqrt{u_{rA1}^2 + u_{rB1}^2} \tag{13a, b}$$

Then their correlation coefficient $\rho_{x1,2}$ is

$$\rho_{x1,2} = \frac{\rho_A u_{rA1} u_{rA1} + \rho_B u_{rB1} u_{rB1}}{u_{rx1} u_{rx2}} \tag{13}$$

In general case for types A and B of relative uncertainties u_{rA} and u_{rB}, in a similar way as before for u_A and u_B, a matrix U_{rX} for components of relative uncertainties can be created. If the relative uncertainties of the input quantities are known, e.g. the same for the entire range, you can directly use their matrix propagation equation with a similar structure as for absolute uncertainties, i.e.:

$$\boldsymbol{U}_{rY} = \boldsymbol{S}_r \cdot \boldsymbol{U}_{rX} \cdot \boldsymbol{S}_r^{\mathrm{T}} \tag{14}$$

where: $\boldsymbol{U}_{rX}$, $\boldsymbol{U}_{rY}$,$\boldsymbol{S}_r = \frac{x_i}{y_j}\frac{\partial y_j}{\partial x_i}$ - covariance matrices for relative uncertainties and the sensitivity matrix with elements marked with indices $i = 1, 2$ for lines and
j = 1, 2. for columns.

5 The components of the uncertainty of the output quantities

When instruments and measuring systems in various environmental conditions are used and measurement observations are different randomly distributed, it may also be necessary to find the type A and B uncertainties of elements of the output Y measurand and correlations between the pair of their values. We analyze this on the example of 2D measurand based on (7a) and (7b) equations. The covariance matrix of measurand $\boldsymbol{X} = [x_1, x_2]^T$ is

$$\boldsymbol{U_X} = \begin{bmatrix} u_{x1}^2 & u_{x12}^2 \\ u_{x12}^2 & u_{x2}^2 \end{bmatrix} = \begin{bmatrix} u_{1A}^2 + u_{1B}^2 & \rho_A u_{1A} u_{2A} + \rho_B u_{1B} u_{2B} \\ \rho_A u_{1A} u_{2A} + \rho_B u_{1B} u_{2B} & u_{2A}^2 + u_{2B}^2 \end{bmatrix} \tag{15}$$

It can be presented as the sum of two matrices for uncertainty of type A and Type B components, ie.

$$\boldsymbol{U}_X = \boldsymbol{U}_{XA} + \boldsymbol{U}_{XB} \tag{16}$$

in which: $\boldsymbol{U_{XA}} = \begin{bmatrix} u_{1A}^2 & \rho_A u_{1A} u_{2A} \\ \rho_A u_{1A} u_{2A} & u_{2A}^2 \end{bmatrix}$, $\boldsymbol{U_{XB}} = \begin{bmatrix} u_{1B}^2 & \rho_B u_{1B} u_{2B} \\ \rho_B u_{1B} u_{2B} & u_{2B}^2 \end{bmatrix}$

The transformation of the covariance matrices U_{XA} and U_{XB} of the uncertainties of measurand X is performed after the linearization of the functional $\boldsymbol{Y} = \boldsymbol{F}(\boldsymbol{X})$ according to formula (2) as follows

$$\boldsymbol{U}_{YA} = \boldsymbol{S} \cdot \boldsymbol{U}_{XA} \cdot \boldsymbol{S^T} \qquad \textit{and} \qquad \boldsymbol{U}_{YB} = \boldsymbol{S} \cdot \boldsymbol{U}_{XB} \cdot \boldsymbol{S^T} \tag{17}$$

The output covariance matrix $\boldsymbol{U}_Y$ can be determined in two ways:

1. estimate the matrix $\boldsymbol{U}_X = \boldsymbol{U}_A + \boldsymbol{U}_B$ and determine from it directly $\boldsymbol{U}_Y$ as follows

$$\boldsymbol{U}_Y = \boldsymbol{S} \cdot \boldsymbol{U}_X \cdot \boldsymbol{S^T} = \boldsymbol{S} \cdot (\boldsymbol{U}_{XA} + \boldsymbol{U}_{XB}) \cdot \boldsymbol{S^T} \tag{18}$$

2. or from the matrix $\boldsymbol{U}_{XA}$ and $\boldsymbol{U}_{XB}$ find the matrices $\boldsymbol{U}_{YA}$ and $\boldsymbol{U}_{YB}$ of both components of the uncertainties u_{yi} of output quantities and then $\boldsymbol{U}_Y$. It is received

$$\boldsymbol{U}_Y = \boldsymbol{U}_{YA} + \boldsymbol{U}_{YB} = \boldsymbol{S} \cdot \boldsymbol{U}_{XA} \cdot \boldsymbol{S}^{\boldsymbol{T}} + \boldsymbol{S} \cdot \boldsymbol{U}_{XB} \cdot \boldsymbol{S}^{\boldsymbol{T}} \tag{19}$$

If $\boldsymbol{U}_Y$ is only designated, then both methods are equivalent. This was checked for both two-dimensional (2D) measurands $\boldsymbol{X}$ and $\boldsymbol{Y}$. On the other hand, matrices $\boldsymbol{U}_{YA}$ and $\boldsymbol{U}_{YB}$ for the uncertainties of type A and B components of the output $\boldsymbol{Y}$ measurand and the correlation coefficients of their each type can be determined from the formula (19) not found in the literature.

6 Uncertainty of sum and difference of output quantities

If $\boldsymbol{Z} = \boldsymbol{G}(\boldsymbol{Y}) = [y_1 + y_2, y_1 - y_2]^{\mathrm{T}}$,then absolute uncertainties are equal:

$$u_{z1}^2 = u_{1A}^2 + u_{1B}^2 + u_{2A}^2 + u_{2B}^2 + 2(\rho_A u_{1A} u_{2A} + \rho_B u_{1B} u_{2B}), \tag{20a}$$

$$u_{z2}^2 = u_{1A}^2 + u_{1B}^2 + u_{2A}^2 + u_{2B}^2 - 2(\rho_A u_{1A} u_{2A} + \rho_B u_{1B} u_{2B}) \tag{20b}$$

Let use designation $u_0^2 \equiv u_{1A}^2 + u_{1B}^2 + u_{2A}^2 + u_{2B}^2$ for the output variance of the sum and the difference of values in the absence of correlation.Then

$$u_{y1+y2} = \sqrt{u_0^2 + 2(\rho_A u_{1A} u_{2A} + \rho_B u_{1B} u_{2B})} \tag{21a}$$

$$u_{y1-y2} = \sqrt{u_0^2 - 2(\rho_A u_{1A} u_{2A} + \rho_B u_{1B} u_{2B})} \tag{21b}$$

It is obtained that the uncertainty increases for the sum and decreases for the difference in relation to the model with zero input correlation. The values $y_1 + y_2$ and $y_1 - y_2$ are correlated with a coefficient equal to:

$$\rho_{y1+y2,y1-y2} = \frac{u_{y1}^2 - u_{y2}^2}{u_{y1+y2} u_{y1-y2}} \tag{22}$$

If there is no correlation between type B components, i.e. for $\rho_B = 0$, uncertainty values are determined by

$$u_{y1+y2} = \sqrt{u_0^2 + 2\rho_A u_{1A} u_{2A}} \quad \textit{and} \quad u_{y1-y2} = \sqrt{u_0^2 - 2\rho_A u_{1A} u_{2A}} \tag{23a, b}$$

In this case, we obtain a reduction in the uncertainty for the sum of the input quantities and an increase in the uncertainty for the difference of this quantities. The numeric values is demonstrated in the example of the uncertainty of measurements of the mean value and the difference of two temperatures given in [13] and [16].

Summary

This paper analyzes the uncertainty model of two-parameter measurements, which is a development of the internationally used model according to the guide for evaluation of uncertainty in measurements under acronym GUM [1]. It was assumed that in the general case the uncertainty components of each type A and B for quantities jointly measured in the input of measurement system can be correlated with each other. The general dependences of the correlation coefficient and the different mutual relations between input uncertainties were determined, with full correlation and no correlation between type B uncertainties. The uncertainties of the output quantities and the correlation coefficient for the linear function processing the input quantities and for the quadratic-linear and quadratic functions were determined. quotient. The dependence of the correlation coefficient of output values on the values of input parameters was investigated. The content of the work contains the courses of the dependencies studied and several detailed conclusions.

In cases when both input values are measured in equal or similar influencing conditions, their B-type uncertainties may be correlated. This should be considered in the estimation of the uncertainty of the output quantities. For example, for the sum of the output quantities and the positive correlation coefficient, the result of uncertainty will be greater than the result from the geometric summation of both uncertainties type A and type B according to GUM, and for the difference - smaller. Considerations and conclusions can be generalized to measurements of many multiparameter measurands.

References

[1] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML,Evaluation of measurement data — Guide to the Expression of Uncertainty in Measurement, JCGM 100:2008, GUM 1995 with minor corrections.

[2] Supplement 1 to GUM—Propagation of distributions using a Monte Carlo method. JCGM 101:2008, BIPM.

[3] Supplement 2 to GUM - Extension to any number of output quantities. JCGM 102:2011 BIPM.

[4] Supplement 6 to GUM - The role of measurement uncertainty in conformity assessment JCGM 106:2012, BIPM.

[5] Conformity Assessment - Requirements for the Competence of Testing and Calibration Labs, ISO/IEC 17025:2017.

[6] EA-4/02.M: 2013 Evaluation of the Uncertainty of Measurement in Calibration.

[7] Zakharow I.P. (2007). Estimating measurement uncertainty on the basis of observed and logical correlation *Measurement techniques*. Vol. **50**, No.8.

[8] Dorozhovets M., Warsza Z. L. (2007). Upgrading calculating methods of the uncertainty of measurement results in practice *Przegląd Elektrotechniki -Electrical Review*. Vol. **1**, pp. 1-13(in Polish).

[9] Dorozhovets M., Warsza Z. L. (2007). Proposals of Upgrading the Calculating Methods of the Measurement Result Uncertainty given by GUM. (Part 2) Development of the method type B. *Pomiary Automatyka Robotyka (PAR)*. Vol. **2**, pp. 6-12(in Polish).

[10] Warsza Z.L. Part 1, and Warsza,Z. L., Puchalski. (2018). Part 2: Estimation of uncertainty of indirect measurement in multi-parametric systems with few examples. PPt: in CD Proceedings of conference: Problems and Progress of Metrology ppm'18-Szczyrk 04-06.*Series: Conferences No. 22, Metrology Commission of Katowice Branch of the Polish Academy of Science, ISBN 978-83-7880-541-0 on CD.* (in Polish).

[11] Warsza,Z. L., Puchalski, J. (2018). Part 2: Estimation of vector uncertainties of multivariable indirect instrumental measurement systems on the star circuit example. *XXII World Congress IMEKO 2018 Belfast. Abstract in CD Proceedings PO-062 and IOP Conf. Series: Journal of Physics: Conf. Series 1065 052026*, DOI:10.1088/1742-6596/1065/5/052026.

[12] Warsza Z.L., Idzikowski A. (2018). Accuracy description of circuits with Rtd sensors dedicated to the temperature difference and average measurements. in book: Ed. Świder, J. et all: Mechatronics 2017 - Ideas for Industrial Applications, series: AISC 914, 2018 Springer Nature Switzerland AG. pp.435-446. doi.org 10 1007 978-3-030-15856-9

[13] Warsza Z.L., Puchalski J. (2018). Matrix estimation of uncertainty of indirect multiparameter measurements with examples *Pomiary Automatyka Robotyka. (PAR) 2 2018*. Vol. **2**, pp. 31-40 DOI:1014311 PAR 228 31-40 (in Polish).

[14] Warsza Z.L., Puchalski J. (2018). Evaluation of the uncertainties of rectangular components of the impedance determined indirectly from measurements of polar components and vice versa *Pomiary Automatyka Robotyka. (PAR) 3 2018*. Vol. **3**, pp. 61-67 DOI:10.14313 PAR 229 (in Polish).

[15] Warsza Z.L., Puchalski J. (2018). Estimation of the uncertainty of function values from measurements at control points *Pomiary Automatyka Robotyka. (PAR) 4 2018*. Vol. **4**, pp. 39-50 DOI:10.14313 PAR 230 (in Polish)-presented on AMSAV.

[16] Warsza Z.L., Puchalski J. (2019). Upgraded Method of the Estimation of Uncertainties in Multiparameter Measurements.Part 1.Theoretical Basis for Correlated Measured Variables. *Pomiary Automatyka Robotyka. (PAR) 1 2019*. Vol. **1**, pp. 47-57 DOI: 10.14313 PAR 231 47 (in Polish).

Application of the vector method for estimating characteristic function based on measurements uncertainty at two control points

Zygmunt Lech Warsza[1], Jacek Puchalski[2] and Adam Idzikowski[3]
[1] *Industrial Research Institute of Automation and Measurements (PIAP), Warszawa Poland*
[2] *Central Office of Measures (GUM), Warszawa Poland*
[3]*Białystok University of Technology, Faculty of Electrical Engineering, Białystok Poland*
e-mail: zlw1936@gmail.com,
jacek.puchalski@gum.gov.pl,a.idzkowski@pb.edu.pl

Abstract

This paper discusses two methods for estimating the uncertainty of values of the function which describes a characteristic of the property of tested devices, substances or engineering process. This estimation is based on measurements at two control points. The method I estimates uncertainty of these function points using a linear approximation created on the basis of maximum permissible values of measurement error at control points. The method II relies on the statistical estimation of values, their uncertainties and correlation for the points of a tested function as a linear combination of measurement results in two control points. Matrix approach to the propagation of uncertainties in indirect multivariate measurements was used. Method I is the boundary case of method II when the correlation coefficient is equal to 1. Using the method II, the absolute and relative uncertainties of interpolated values of characteristic curve and their linear or nonlinear functions can be properly estimated. Both methods can be useful in all areas of metrology applications. It is an extension of the method described in the GUM Guide Supplement 2.

Keywords: maximum permissible error, uncertainty type A and type B, correlation coefficient, two-parameter measurand, control points, statistical estimation, multivariate measurements, vector propagation of uncertainty.

Introduction

To assess the accuracy of measurements, the concept of uncertainty was introduced in the 1990s and the principles of its application were published in the form of the Guide to the Expression of the Uncertainty in Measurement (GUM) [1] and its Supplements. The scope of application of these international recommendations has been still expanding in numerous publications, including a monograph [2]. In addition to the widespread use of the concept "uncertainty" to evaluate the accuracy of measurements, new applications for the assessment of product quality, statistical quality control of production processes and laboratory accreditation have emerged and spread in experimental research [2-11].

In experimental studies, there are usually several limitations concerning density of measurement points, time taken and duration of the experiment, availability of a tested object, samples and automated equipment, and cost of performance. In laboratory tests and in only a few routine tests, these are usually not critical requirements. In most utility studies there are such limitations and they are important. It is necessary to minimize the number of controlled points and to select the appropriate distribution, time and volume of the information obtained. It also depends on the type of tests and the range of values of the examined quantities, on the possibility of obtaining the uncertainty of measurements required in these tests and on the accuracy of the measuring equipment used. There may be a requirement that control points should be unevenly distributed even for a linear function for reasons other than measurement accuracy, e.g. consumption of test substances and reagents, volume, dimensional and flow restrictions, power and energy limitations, etc.

The paper considers the cases when estimating values and uncertainties for selected points of the curve modeled with the known function $y = f(x)$, which are not measured directly. It is also necessary to determine the gradients of uncertainty in the analyzed range of this function and the interval x of a given uncertainty. The possibilities of estimation based on control measurements in two points were analyzed. The estimated uncertainties depend on the number and location of these points along the scope of the function being tested. Discussions on this issue could not be found in the literature.

This issue is discussed on the example analysis which aim is to determine uncertainty of the known measurement function. This was the method proposed by the first author, which is based on measuring two values x_1, x_2 only with uncertainties u_{x_1}, u_{x_2} and their correlation estimated from measurement data and experiment conditions. To evaluate the accuracy of measurement results, recommendations included in the GUM Guide [1] were applied. The basis of this assessment is the estimation of standard uncertainty as a geometric sum of component uncertainties i.e. $u = \sqrt{u_A^2 + u_B^2}$. The component u_A, which is called as the type A uncertainty, is determined by the scattering of the specified number n of repeated measurements of the measured quantity in circumstances considered as random ones. The u_B component, i.e. the type B uncertainty, represents the randomized cumulative impact of the predicted various impacts on test object, measurement system and instrument readings when using them under the specified permissible conditions and in the nominal lifetime [1,10,11]. Interferences that occur during measurements can cause systematic errors of unknown values. Basing on the knowledge about predicted ranges and probability distributions of different interferences, heuristically their contributions in u_B uncertainty are estimated and statistically for the long time this uncertainty component type B is determined [5,6].

Based on the measured control values x_1, x_2 and their standard uncertainties u_1, u_2 the value of x_{ci} and its uncertainties $u_{ci} \equiv \sigma_{ci}$ were estimated. The dependencies for indirect estimation of the absolute standard deviation $\sigma_{xi)}$ and the relative deviation $\delta_i \equiv \sigma_{xi} \cdot x_i$ were determined. These estimates make possible to determine the ranges of x with given end-of-range values as well as uncertainties of quantities depending

on x with the known function.

Two methods of estimating values and uncertainties for any number of function points describing the tested characteristic curve will be applied. Method I is based on the determination of standard uncertainty from the permissible maximum error of the instrument, known from reliable calibration results or from data given by the manufacturer. Method II, statistical, is based directly on the results of measurements at control points and on the knowledge about measuring devices and experiment conditions.

1 Method I

In method I, for the measured values of the function of tested object, the type B uncertainty is estimated mainly from accuracy data of measuring instruments [2-6]. The permissible maximal error of instrument $|\Delta_x|_{max}$ is usually specified by manufacturers as the line dependence of the absolute error module on the value x of measured quantity in the form of sum of components, additive and multiplicative or after their relation to the measuring range as a maximum relative error. This is described in the formulas (1) and (1a)

$$|\Delta_x| \leq |\Delta_x|_{max} = |\Delta_{x_0}|_{max} + (x - x_0) \cdot |\varepsilon_S|_{max} \tag{1}$$

$$\frac{|\Delta_x|}{x_{max} - x_0} \leq \frac{|\Delta_{x_0}|_{max}}{x_{max} - x_0} + \frac{x - x_0}{x_{max} - x_0} \cdot |\varepsilon_S|_{max} \tag{1a}$$

where: in (1): $|\Delta_x|$, $|\Delta_x|_{max}, |\Delta_{x_0}|_{max}$- the absolute errors: real of the measured value x and maximum permissible for x and for the beginning of range x_0, e.g. $x_0 = 0$ and $|\varepsilon_S|_{max} \equiv \Delta_{x-x_0}/(x-x_0)$ - maximum permissible relative error of the difference $x-x_0$. Equation (1a) gives their values related to the range and is simplified for $x_0 = 0$. With the method I, from two maximum values of this error $|\Delta_{x_1}|_{max}, |\Delta_{x_2}|_{max}$, which are known from technical data or from measurements under given conditions, for the x value in the tested range one can determine the linear characteristic curve of absolute uncertainty σ_x or its normalized value (the relative uncertainty). Within the ranges $\pm|\Delta_{x_1}|_{max}$ $\pm|\Delta_{x_2}|_{max}$, the dispersion of the x-value with a uniform probability distribution is usually assumed. Standard deviations σ_{x1}, σ_{x2} of estimators of the values x_1, x_2, e.g. mean values of measurement observations, are treated as absolute type A uncertainties. In the field of uncertainty, the formula (1) corresponds to relationship given below

$$\sigma_x \leq \sigma_x|_{max} = \sigma_{x0}|_{max} + (x - x_0) \cdot \delta_x|_{max} \tag{2}$$

where: $\delta_x|_{max}$ is maximum relative uncertainty of the difference $(x - x_0)$.
The maximum permissible absolute error $|\Delta_x|_{max}$ from equation (1) and the corresponding absolute uncertainty $\sigma_x|_{max}$ from equation (2) are proportional to each other $(1/\sqrt{3})$. With appropriately selected scales on the y-axis, they run in the function of the measured quantity x identically linear, as shown in Figure 1.

Method I is very simple. It can be used to the approximate estimation of uncertainty in the full range or only in those parts of it, i.e. outside the interval (x_1, x_2), where it can be assumed the linear dependence of the maximum permissible error on x. For example, for many digital devices, the maximum permissible absolute error increases from z $x_2 \geq x_1$, i.e. $|\Delta_{x_2}|_{max} \geq |\Delta_{x_1}|_{max}$. If the relative error $|\delta|_x$ is constant, then this increase in type B uncertainty is proportional to the increment of x.

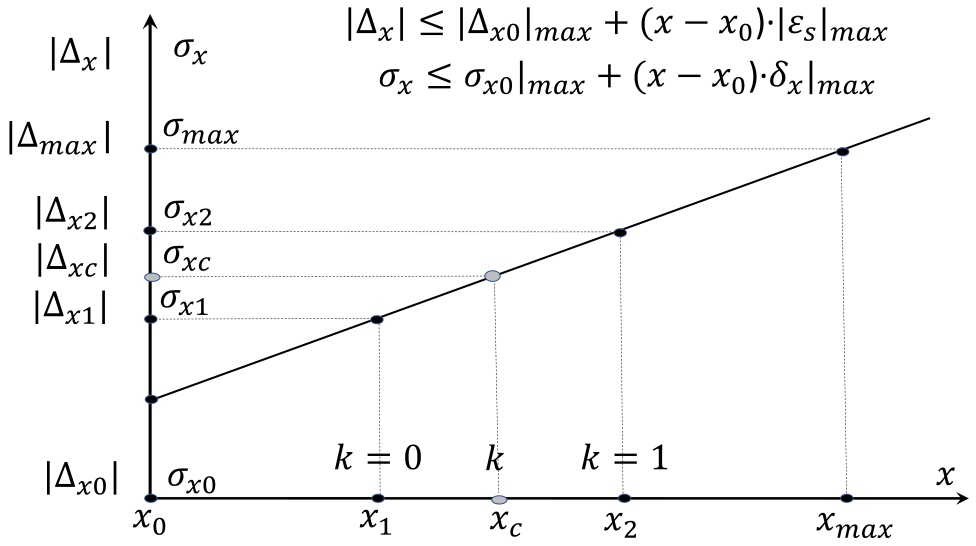


Figure 1: Linear characteristic curve of maximum permissible absolute error $|\Delta_x|_{max}$ given in formula (1) and corresponding absolute uncertainty $\sigma_x|_{max}$ in the $\sqrt{3}$ larger scales on the y-axis

However, when using the method I to estimate uncertainty, for individual x_c values one cannot take into account the statistical nature of the uncertainty of controlled values x_1, x_2, including the correlation of type B uncertainty components of the measuring instruments. These components are estimated considering the maximum permissible errors of instruments, and the correlation coefficient between the obtained uncertainties for any values within the measuring range is equal to 1. The correlation coefficients between estimated uncertainties for x_{c1},x_{c2} cannot be determined and the correlation coefficient of 1 is assumed for them. When determining the uncertainty of sum and difference of two values x with the correlation coefficient equal to $\pm$ 1, the component uncertainties should be added algebraically. For the sum, they are larger, and for the difference - smaller than for the geometric summation.

Therefore, the application of method I is limited for determining uncertainty of indirect multivariate measurements. It is only suitable for these cases, when the individual x_c values of the characteristic curve with estimated uncertainties will be used later only individually.

2 Method II - statistical

Method II relies on a statistical description of accuracy using uncertainty. It is based on the vector method for determining uncertainty of multivariate indirect measurements, which is presented in Supplement 2 to the GUM Guide [1]. The output $\boldsymbol{Y}$ multi-measurand parameters are determined indirectly from the $\boldsymbol{X}$ multi-measurand input measurements. The general function is

$$\mathbf{Y} = F(X). \tag{3}$$

Their uncertainties and correlation coefficients are related to the Law of Propagation of Variances as a relation of the covariance matrices U_Y and U_X,i.e.:

$$U_Y = S \cdot U_X \cdot S^T \tag{4}$$

If the relative uncertainties do not exceed several percent, then from the equation (4) it is determined the uncertainty for processing function$Y = F(X)$, both linear and nonlinear, for which the sensitivity matrix $S = \frac{\partial Y}{\partial X}$ is Jacobian matrix. For nonlinear functions, the dimension m of the vector Y may be larger than the dimension n of the vector X, equal to the number of independent equations connecting elements of both vectors.

The issues discussed should be general and useful for any type of $F()$ function. Therefore, the uncertainty evaluation of the elements Y is divided into two stages. In the first of them, based on measurement results of two values x_1, x_2, a linear scale of values $X_c = F_c(X)$ is created for the full range of $x_{max} - x_0$ considered and estimates their uncertainties. In the second stage, from the selected elements x_c of the vector X_c, the values and uncertainties of the elements of the vector Y are determined according to the individual linear or non-linear function $Y = F_y(X_c)$. The first stage will be discussed wider. The uncertainty of any x_c value results from the control measurements x_1, x_2, their uncertainty and the correlation coefficient. It is their linear combination described by formula (5)

$$x_c = x_1 + k\,(x_2 - x_1) = (1 - k)\,x_1 + kx_2 \tag{5}$$

where k means relative location of the point x_c in the interval x_1, x_2.

$$k = (x_c - x_1)/(x_2 - x_1) \tag{5a}$$

Inside the interval $x_1 \le x_c \le x_2$ and k is $0 \le k \le 1$.

The values obtained in control measurements x_1, x_2, estimated values of x_c and uncertainties $\sigma_{x1}, \sigma_{x2}, \sigma_c$ are modeled with random variables. The uncertainties σ_c are evaluated indirectly from uncertainties σ_{x1}, σ_{x2} and their correlation coefficient $\rho_{x1,2}$.

As an example, the measurand X_c of two values x_{c1}, x_{c2} will be computed . Their uncertainties σ_{c1}, σ_{c2} and correlation coefficient $\rho_{c1;;2}$ will be evaluated. When estimating these parameters of the X_c its covariance matrix U_c for absolute uncertainties is as follows

$$U_c = S_c \cdot U_X \cdot {S_c}^T \tag{6}$$

where: sensitivity matrix S_c and covariance matrices U_X and U_c of vectors $X = [x_1, \ x_2]^{\mathrm{T}}$ and $X_c = [x_{c1}, x_{c2}]^T$ are described by equations (6a, b, c) given below

$$S_c = \begin{bmatrix} \frac{\partial x_{c1}}{\partial x_1} & \frac{\partial x_{c1}}{\partial x_2} \\ \frac{\partial x_{c2}}{\partial x_1} & \frac{\partial x_{c2}}{\partial x_2} \end{bmatrix}, \quad U_X = \begin{bmatrix} \sigma_{x1}^2 & \rho_{x1,2}\sigma_{x1}\sigma_{x2} \\ \rho_{x1,2}\sigma_{x1}\sigma_{x2} & \sigma_{x2}^2 \end{bmatrix} \tag{6a, b}$$

$$U_c = \begin{bmatrix} \sigma_{c1}^2 & \rho_{c1,2}\sigma_{c1}\sigma_{c2} \\ \rho_{c1,2}\sigma_{c1}\sigma_{c2} & \sigma_{c2}^2 \end{bmatrix} \tag{6c}$$

where: σ_{x1}, σ_{x2}, σ_{c1},σ_{c2} are absolute uncertainties and $\rho_{x1,2}$, $\rho_{c1,2}$ correlation coefficients.

If two calculated values x_{c1}, x_{c2} are linear with the measured values x_1, x_2 , then they are represented by linear combinations and defined by equations

$$x_{c1} = (1 - k_1)\, x_1 + k_1 x_2 \tag{7a}$$

$$x_{c2} = (1 - k_2)\, x_1 + k_2 x_2 \tag{7b}$$

In the general case, the results of control measurements x_1, x_2 are correlated, i.e. $\rho_{x1,2} \neq 0$. The initial covariance matrix U_c is derived from (6) and (7a,b) and it is

$$U_c = \begin{bmatrix} 1-k_1 & k_1 \\ 1-k_2 & k_2 \end{bmatrix} \cdot \ U_X \ \cdot \ \begin{bmatrix} 1-k_1 & 1-k_2 \\ k_1 & k_2 \end{bmatrix} \tag{8}$$

The main diagonal of the matrix U_c consists of the elements which are the variances of values x_{c1} and x_{c2}, i.e. the squares of uncertainties $\sigma_{xc1}, \sigma_{xc2}$ given by equations

$$\sigma^2{}_{c1} = (1 - k_1)^2 \sigma^2{}_{x1} + k_1{}^2 \sigma^2{}_{x2} + 2\rho_{x1,2}(1 - k_1)k_1\sigma_{x1}\sigma_{x2} \tag{8a}$$

$$\sigma^2{}_{c2} = (1 - k_2)^2 \sigma^2{}_{x1} + k_2{}^2 \sigma^2{}_{x2} + 2\rho_{x1,2}(1 - k_2)k_2\sigma_{x1}\sigma_{x2} \tag{8b}$$

Their normalization to uncertainty σ_{x2} of control point x_2 is given by the formula (9) where:i=1, 2, and $\varepsilon = \frac{\sigma_{x1}}{\sigma_{x2}}$ is the ratio of uncertainties for x_1 and x_2.

$$\sigma_{nci} = \frac{\sigma_{ci}}{\sigma_{x2}} = \sqrt{\varepsilon^2(1 - k_i)^2 + k_i{}^2 + 2\rho_{x1,2}\varepsilon(1 - k_i)k_i} \tag{9}$$

Correlation coefficient $\rho_{c1,2}$ of the estimated quantities x_{c1}, x_{c2} is given by the formula

$$\rho_{c1,2} = \frac{\varepsilon^2(1 - k_1)(1 - k_2) + k_2k_1 + (k_1 + k_2 - 2k_1k_2)\varepsilon\rho_{x1,2}}{\sigma_{nc1}\sigma_{nc2}} \tag{10}$$

From (10), for special cases: of uncorrelated x_1 and x_2, ($\rho_{x1,2} = 0$) and x_1 and x_2 perfectly correlated ($\rho_{x1,2} = 1$) derived are formulas (11) and (12):

$$\rho_{c1,2} = \frac{\varepsilon^2(1 - k_1)(1 - k_2) + k_1k_2}{\sigma_{nc1}\sigma_{nc2}} \tag{11}$$

$$\rho_{c1,2} = 1 \tag{12}$$

The dependence of uncertainty σ_c on k of point x_c is described by the formula (13) and its special cases - formulas (14) and (15)

$$\sigma_c = \sqrt{(1-k)^2\sigma^2{}_{x1} + k^2\sigma^2{}_{x2} + 2k(1-k)\rho_{x1,2}\sigma_{x1}\sigma_{x2}} \tag{13}$$

If is no correlation $\rho_{x1,2} = 0$, or perfect correlation $\rho_{x1,2} = 1$ (13) is simplified as follows

$$\sigma_c|_{\rho_{x1,2}=0} = \sqrt{(1-k)^2\sigma^2{}_{x1} + k^2\sigma^2{}_{x2}} \tag{14}$$

$$\sigma_c|_{\rho_{x1,2}=1} = (1-k)\,\sigma_{x1} + k\sigma_{x2} \tag{15}$$

The formula (15) becomes linear.The correlation coefficient $\rho_{c1,2}$ is the function of relative locations k_1, k_2 of the estimated values x_{c1}, x_{c2}, for the ratio of uncertainties $\varepsilon = 1/2$ of control points x_1, x_2 and their correlation coefficient $\rho_{x1,2} = (0; 0.5; 1)$ is presented in Fig. 2.

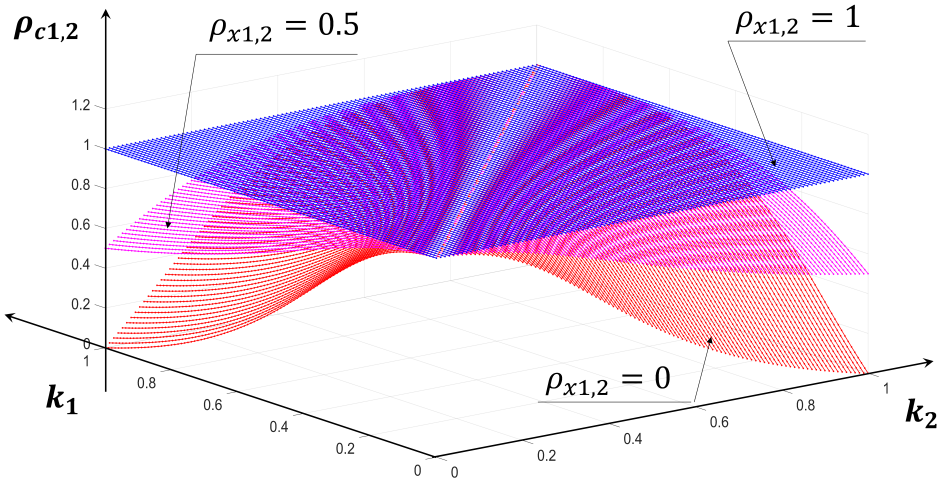


Figure 2: Correlations coefficient $\rho_{c1;2}$ as a function of the relative locations k_1, k_2 of values x_{c1}, x_{c2} for their correlation coefficients $\rho_{x1,2} = (0; 0.5; 1)$ and the ratio of uncertainties $\varepsilon = 1/2$

Figure 2 shows that 2D surfaces of the correlation coefficient $\rho_{c1,2} = f(k_1, k_2)$ for $\rho_{x1,2} < 1$ on the input reach a maximum of 1 for $k_1 = k_2$. The coefficient $\rho_{c1,2} = 1$ creates a plane in the entire area of variability k_1, k_2. With decreasing $\rho_{x1,2}$, the symmetric surfaces $\rho_{x1,2}$ against the line $k_1 = k_2$ decrease and reach the minimum in symmetrically located points $k_1 = 1, k_2 = 0$ and $k_1 = 0, k_2 = 1$

After normalizing σ_c and σ_{x1} to σ_{x2}i.e. for $\varepsilon = \frac{\sigma_{x1}}{\sigma_{x2}}$ equation(16)-(18) are derived

$$\frac{\sigma_c}{\sigma_{x2}} = \sqrt{\varepsilon^2(1-k)^2 + k^2 + 2\rho_{x1,2}\varepsilon(1-k)k} \tag{16}$$

for $\rho_{x1,2} = 0$ $$\frac{\sigma_c}{\sigma_{x2}} = \sqrt{\varepsilon^2(1-k)^2 + k^2} \quad (17)$$

for $\rho_{x1,2} = 1$ $$\frac{\sigma_c}{\sigma_{x2}} = \sqrt{(\varepsilon - k\varepsilon + k)^2} = \varepsilon + k(1-\varepsilon) \quad (18)$$

From (10) and (16) the correlation coefficient of $x_{c1} = x_1$ and $x_{c2} = x_c$ (i.e.$k_1 = 0$ and $k_2 = k$) is

$$\rho_{x_1,c} = \frac{\varepsilon^2(1-k)}{\sqrt{\varepsilon^2(1-k)^2 + k^2}} \quad (19)$$

Resulting from (19) figure 3 gives the dependence of the correlation coefficient $\rho_{x_1,c}$=$f(k)$ for uncertainty ratios ε=(0.5; 2/3; 1) of control points for the input correlation coefficient $\rho_{x1,2} = 0$. With the increase of k to 1, all curves reach 0, for lower ε strongly.

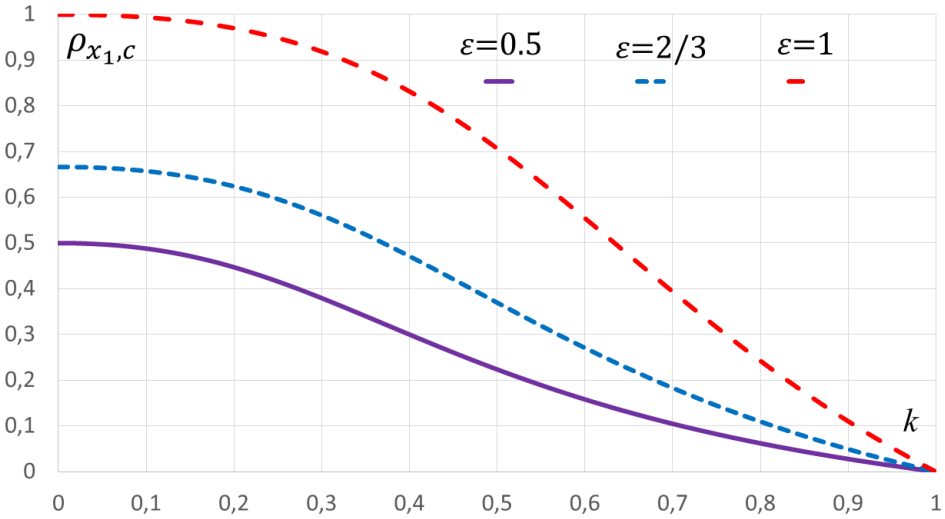


Figure 3: Correlation coefficient ρ_{x_1c} as a function of the relative location k of point x_c inside of the interval x_1, x_2 for few ratios ε=(1/2; 2/3; 1) of uncertainties for its ends

The values x_{c1} and x_{c2} can also be used to model the multivariate measurand $\boldsymbol{Y}$ determined by the vector function $Y = F_y(X_c)$.This function can be both linear and non-linear. For example, for both two-element vectors $\boldsymbol{Y} = [y_1,\ y_2]^T$, X_c= $[x_{c1},\ x_{c2}]^T$ their covariance matrices U_Y i U_c are related according to the Law of Propagation of variances (4), (20)

$$U_Y = S_Y \cdot U_X \cdot {S_Y}^T \quad (20)$$

where $S_Y = \begin{bmatrix} \frac{\partial y_1}{\partial x_{c1}} & \frac{\partial y_1}{\partial x_{c2}} \\ \frac{\partial y_2}{\partial x_{c1}} & \frac{\partial y_2}{\partial x_{c2}} \end{bmatrix}$ Uncertainties and the correlation coefficient of output quantities y_1, y_2 are determined here from U_Y matrix. At first, the input values

x_{c1} and x_{c2} and their covariance matrix U_c are obtained. Next, formulas for elements of the covariance matrix U_Y are used. The value of x_c, the uncertainty of which is estimated in the linear case, is related to t values x_1, x_2 by the same formula (3) as for the interpolation using method I. The linear dependence (3) also applies outside the endpoints of extrapolation interval x_c, i.e. if $k \leq 0$ and $k \geq 1$. Generally, for correlated x_1, x_2 it results from (9a, b) a square of normalized uncertainty which is extrapolated in the full range of k. It is described by the formula (21)

$$\sigma_{nc}^2 = \frac{\sigma_c^2}{\sigma_{x2}^2} = \varepsilon^2 (1-k)^2 + k^2 + 2(1-k)k\rho_{x1,2} \tag{21}$$

The dependence of the normalized uncertainty $\sigma_{nc} = f(k, \rho_{x1,2})$ of the point x_c from its relative location k and the correlation coefficient $\rho_{x1,2}$ for three uncertainty ratios ε of points x_1, x_2 are plotted on the 3D graph in Fig. 4.

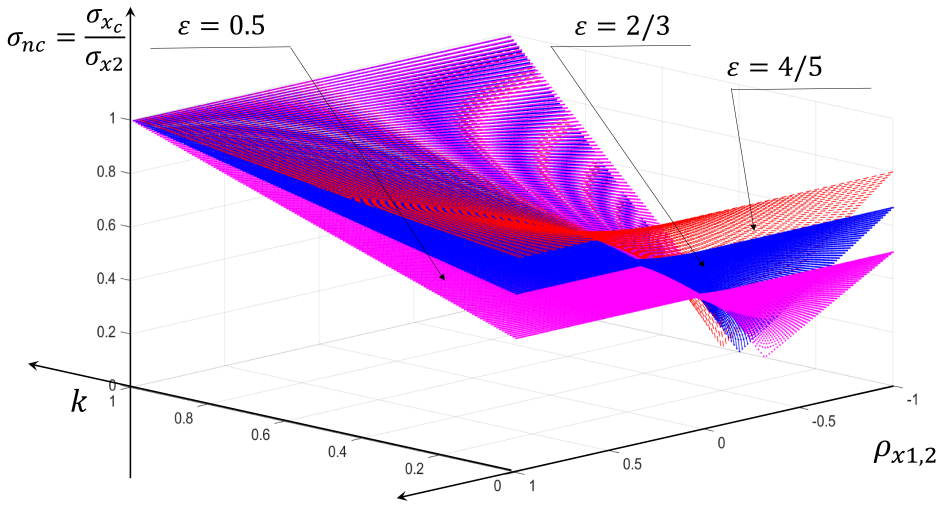


Figure 4: The normalized uncertainty of value x_c as a function of its relative location k and correlation coefficient $\rho_{x1,2}$ for three ratios of uncertainty of input quantities $\varepsilon = \frac{\sigma_{x1}}{\sigma_{x2}} = 0.5; 2/3; 4/5$

The examples of normalized uncertainty graphing σ_{nc} according to method I and formula (2) – the straight line ($\rho_{x1,2} = 1$) and according to method II - surface cross-sections in Fig. 4 for $\rho_{x1,2} = 0$ and two uncertainty ratios ε are presented in Fig. 5. and σ_{nx}=σ_{nc} for $\rho_{x1,2} = 1$.

The uncertainty σ_{nc} rises with k and $\rho_{x1,2}$. From differentiation of the function (21) with respect to k it results that inside the interval $< x_1, x_2 >$ for the value k_{min} given

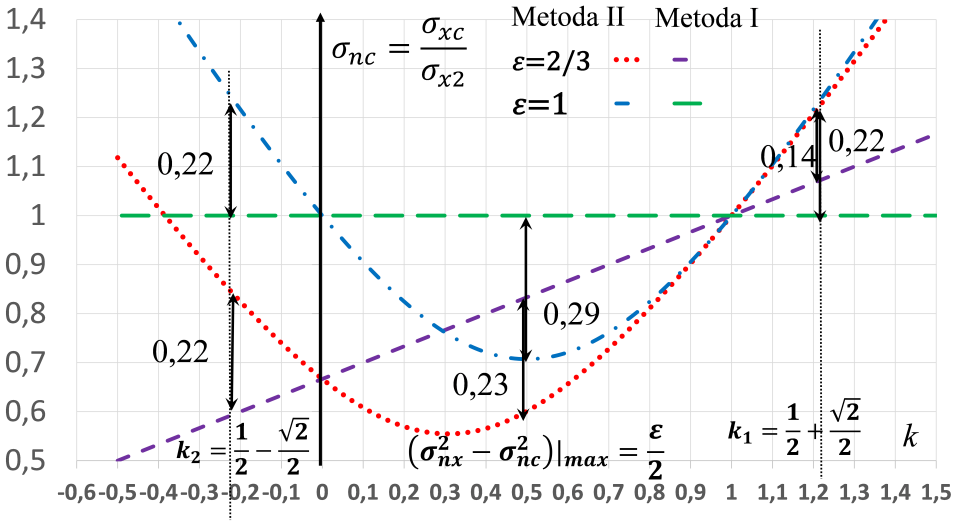


Figure 5: Normalized to σ_{x2} absolute uncertainties of value x_c as a function of its related location k for ratios of uncertainties $\varepsilon = (2/3; 1)$ $\rho_{x1,2} = (1; 0)$; (methods I and II)

in the formula (22) the uncertainty σ_{nc} reaches the minimum

$$k_{min}\,(\sigma_{nc} = min) = \frac{\varepsilon(\varepsilon - \rho_{x1,2})}{1 + \varepsilon^2 - 2\varepsilon\rho_{x1,2}} \tag{22}$$

From (22) it results that relative location k_{min} of point with minimum of the uncertainty σ_{nc} in the control interval x_1, x_2 it does not depend on its width $x_2 - x_1$, but on the relationship of ratio of uncertainties ε and $\rho_{x1,2}$. The absolute location of minimum is defined as

$$x_c|_{\sigma_{nc}=min} = x_1 + k_{min}(x_2 - x_1) \tag{23}$$

The condition $x_1 < x_c|_{\sigma_{nc}=min} < x_2$ is also fulfilled. If there is no uncertainty of one of the endpoints of interval $< x_1,\ x_2 >$, e.g $.\sigma_{x2} \to 0$, then minimum decreases to zero ($k_{\min} \to 0$). The location of the control interval in the range of measurements $x_max - x_0$ does not matter. Fig. 6 shows the dependences of the relative location of $k_{min}(\varepsilon)$ for three values of the correlation coefficient $\rho_{x1,2} = -0.9; 0; +0.9$. They were obtained from (22).

The minimum value of uncertainty is

$$\sigma_{n\ min} = \varepsilon\sqrt{\frac{1 - \rho_{x1,2}^2}{\varepsilon^2 + 1 - 2\rho_{x1,2}\varepsilon}} \tag{24}$$

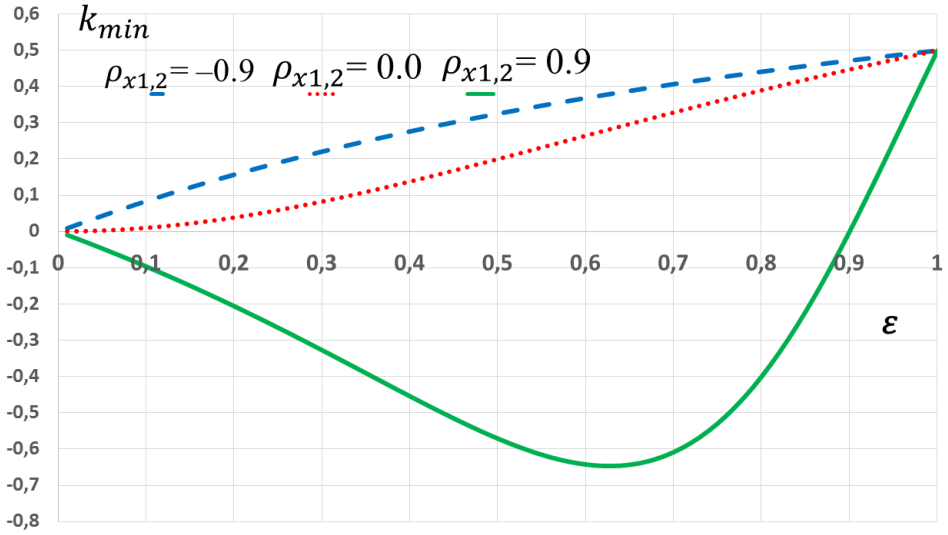


Figure 6: The location k_{min} of minimum uncertainty (method II) as functions of ε

The function $\sigma_{n\ min} = f(\varepsilon)$ are presented in Fig. 7.

And it shows that uncertainty estimated using method II $\sigma_{n\ min}$ in the interval x_1, x_2 is lower about 30% from the uncertainty at the starting point x_1 for $\varepsilon = 1$. For $\varepsilon = 0.6$ this reduction does not exceed 30%, and for $\varepsilon \approx 0.8$ the uncertainty is lower from σ_{x1} about 10% When there is no correlation between the control values x_1, x_2, i.e. for $\rho_{x1,2} = 0$, equation (24) is simplified to the form (24a) and minimum uncertainty depends only on the ratio of uncertainties ε of the control points.

$$\sigma_{n\ min} = \frac{\varepsilon}{\sqrt{1+\varepsilon^2}} \tag{24a}$$

The equation (21) was also used for extrapolating uncertainty of points outside the interval x_1, x_2 i.e. for $k < 0$ and $k > 1$. The difference of squares of uncertainty interpolated by both methods at the point $k = \frac{1}{2}$ is

$$(\sigma_{nx}^2 - \sigma_{nc}^2)|_{max} = \frac{\varepsilon}{2}\,(1 - \rho_{x1,2}) \tag{25}$$

Dependence of uncertainty on the relative position k for method II is parabolic. When extrapolating by using formula (21) the square of uncertainty for $\rho_{x1,2} = 0$ increases by $\frac{\varepsilon}{2}$ at the points $k_{1,2}$ equally distant from the vertex of a parabola with a coordinate k=1/2 as

$$k_{1,2} = \frac{1}{2} \pm \frac{\sqrt{2}}{2} \tag{26}$$

The uncertainty estimated by the second method (method II) is greater at these points than by the first method (method I).

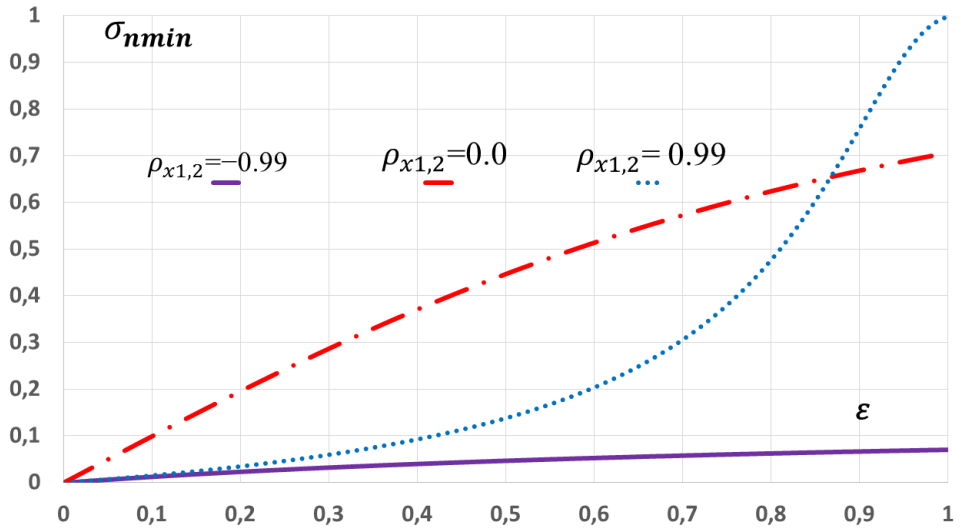


Figure 7: The normalized minimum absolute uncertainty in the interpolation interval as a function of ratio of control point x_1, x_2 for different correlation coefficient $\rho_{x1,2} = (-0.99; 0.00; 0.99)$

Summary and Conclusions

Presented method is proposed by the first author. It is based only on the results of measurements in two points with the required accuracy and it is a complementary to the regression methods. The uncertainties of these points were estimated on the example of a linear function. Correlation coefficients estimated with uncertainties were also determined. This is a quite frequent case, as it concerns measuring devices and systems with direct reading of the measured quantity and multiplier range change. Their resultant function - the output is a straight line with an inclination of 1. The uncertainty for nonlinear functions can be obtained after another transformation by the corresponding linear relationship between the covariance matrices.

The content of this work was connected with several previous publications given in bibliography of [11], in which the authors discussed various examples of estimating the uncertainty of functions in the indirect multivariate measurement electrical circuits.They concerned measurements of associated temperatures, magnetic field induction, systems with several resistances (star circuit, bridge circuit) of direct current (DC) and measurements of impedance components at alternating current (AC). This paper considers two methods for estimating the uncertainty of the function values, deterministic and statistical, designated as method I and method II. In method I, a linear relationship was established between the absolute uncertainty and the value of the measured quantity obtained from the limiting error given for the measuring device. It is a method of accuracy evaluation used by device manufacturers. In this method, a correlation coefficient equal to 1 should be assumed for values in the full measuring range. In method II, a statistical model was adopted for estimated and

controlled values. The values at the points of the examined function were interpolated as a linear superposition of the values measured at the control points. From the uncertainty and the correlation coefficient for these points, absolute and relative uncertainties were estimated at any points of the analyzed function. Parabolic trajectories of uncertainty with the occurrence of its minimum within the interval between control points and a significant increase of uncertainty beyond its extremes were obtained. The relative minimum location in the control interval depends on the ratio of uncertainty and the correlation coefficient of its bounds. It does not depend on the location of this interval in the range.

The parameters of control points were normalized, and the results obtained with both methods and their applicability were compared. If the uncertainties of neighboring measurement points differ at least in the same way as the values of the examined quantity, e.g. for the ratio of these uncertainties $\varepsilon = 0.8$, the difference in uncertainty normalized to the interval bound is 0.1, and for $\varepsilon = 0.9$ it is around 0.01. Estimation of the uncertainty of non-linear processing of the tested function can be determined by the same vector method.

Method I is the special case of method II, when the values at the control points have a correlation coefficient equal to 1. The authors were able to use method II to estimate uncertainty for nonlinear functions in paper [10]. The minimum number of measurement points $n> 2$ for the scope of the function being tested depends not only on itself, but also on its uncertainty, usually drawn as a uniform or linear bar in the distance from this function. It also depends on the obtained statistical parameters of the control measurement results. Based on gained knowledge or after obtaining it in other studies of similar objects the predicted or required simplified function describing uncertainty should be assumed.

For the wide range of non-linear characteristic curve, in practice, it may be useful to divide the entire range of the tested functions into adjacent subranges. If their uncertainty does not increase more than three times ($\varepsilon = 0.7, 0.8$) then the relative interpolation differences of both methods will not exceed 0.2 and 0.12 [12].

As an example of applying the general formulas, the uncertainty of sum and difference of two arbitrary values of the studied function, which are estimated from measurements at the control points, can be determined using the statistical method II. As the example are uncertainties of average and difference of temperatures obtained from measurements by two Pt temperature sensors [9]. If such measurements are made by the same sensor and measuring system and in the same environments and if uncertainty type B is dominated (correlation coefficient is nearly 1), then uncertainties from both measurements must be calculated algebraically but not geometrically as it is without correlation.

References

[1] JCGM 102 (2011)., Evaluation of measurement data – Guide to the expression of uncertainty in measurement. Supplement 2 The extension to any number of output quantities. *BIPM*.

[2] Warsza, Z.L., (2016). Methods of Extension of the Measurement Uncertainty Analysis *Monograph PIAP Warszawa* (Metody rozszerzenia analizy niepewności pomiarów)

[3] Volodarsky, E., Warsza, Z., Kosheva, L., Klevtsova, M. (2019). Uncertainty of Measurement and Reliability of Decision Making on Compliance.Szewczyk, R. et all (eds.) Automation 2019 *AISC 920, Springer Int. Publ. AG 2020,* pp 672-683

[4] ISO 21748 (2010). Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation.*ISO*, Geneva Switzerland

[5] Dorozhovets M., Warsza Z. L. (2007) Upgrading calculating methods of the uncertainty of measurement results in practice. *Przegląd Elektrotechniki 1/2007.*, pp. 1-13 (in Polish).

[6] Zakharow I.P. (2007). Estimating measurement uncertainty on the basis of observed and logical correlation *Measurement techniques.* Vol. **50**, No.8. pp 808-816

[7] Warsza Z. L. (2012). Evaluation and Numerical Presentation of the Results of Indirect Multivariate Measurements. in book ed. by Pavese at all: *Advanced Mathematical & Advances in Mathematics for Applied Sciences, 84, World Scientific Publishing Co. 2012,New Jersey · London · Singapore* , pp. 418-425.

[8] Warsza Z.L. (2018). Estimation of uncertainty of indirect measurement in multiparametric systems with few examples. Part 1, and Warsza,Z. L., Puchalski. Part 2: Mat. Konf.(CD): Problems and Progress of Metrology ppm'18 *Series: Conferences No. 22, Metrology Commission of Katowice Branch of the Polish Academy of Science.*

[9] Warsza, Z.L., Idzikowski, A. 2017. Accuracy description of circuits with RTD sensors dedicated to the temperature difference and average measurements. In book: Ed. Świder, *J. et al.: Mechatronics 2017 - Ideas for Industrial Applications Aseries: AISC 914, 2018 Springer Nature Switzerland AG* pp. 435-446, doi.org/10 1007/978-3-030-15856-9

[10] Warsza Z.L., Puchalski J. (2019). Uncertainty estimation of the data of characteristic from measurements at control points. *Pomiary Automatyka Robotyka. R.22,4/2018*, pp. 39-50 DOI:1014311/PAR_230/399 9 (in Polish)

[11] Warsza Z.L., Puchalski J. (2019). Method of the Estimation of Uncertainties in Multiparameter Measurements of Correlated Quantities. *Proceedings of AMSA V p.*

On estimation algorithms in nonparametric analysis of the current status right-censored data

Irina Yu. Malova[1,3] and Sergey V. Malov[2,3]

[1] *Higher School of Technology and Energy, St. Petersburg State University of Industrial Technologies and Design, St. Petersburg, Russia*

[2] *Theodosius Dobzhansky Center for Genome Bioinformatics, St. Petersburg State University, St. Petersburg, Russia*

[3] *Department of Mathematics, St. Petersburg Electrotechnical University (LETI), St. Petersburg, Russia*

e-mail: malovs@sm14820.spb.edu

Abstract

We consider nonparametric estimation algorithms for current status right-censored data model. In the model right-censored event times are not observed exactly, but at some inspection times. The model covers right-censored data, current status data and life table survival data with a single inspection time. We consider the nonparametric estimation algorithms to obtain three nonparametric estimators for the survival function of failure time: maximum likelihood, pseudo maximum likelihood and the naïve estimator. We discuss large sample properties of the estimators. Using the standard R packages we perform simulations, which compare the estimators under small and moderate sample sizes.

Keywords: survival data, right censoring, interval censoring, current status data, nonparametric estimation.

Introduction

Right-censored survival data model is widely applicable in practice in spite of in many cases the event times (failure or censoring) are not observed exactly, and the investigator observes time interval containing a failure time for each of not missed at follow up individuals having symptoms of disease at the endpoint. In the current status right-censored data model the event is observed in a random inspection time if it occurs before the inspection time or not observed otherwise.

Let T and U be the independent failure and censoring times respectively. Right-censored observation consists of the event time $X = T \wedge U$ and the indicator $\delta = \mathbb{1}_{\{T \le U\}}$. The current status right-censored observation is given as $(W, \kappa, \kappa\delta)$, where $\kappa = \mathbb{1}_{\{X \le W\}}$ and W is a random inspection time, which is independent of (T, U). The observed data is a sample from the distribution $(W, \kappa, \kappa\delta)$ and the main target of statistical analysis is the distribution function F of failure time T.

The right-censored survival data model is well developed. The Kaplan–Meier [19] estimator is widely applicable to estimate the survival function of failure time from right-censored data. Consistency and asymptotic normality of the Kaplan–Meier estimator are obtained first in [5]. The point process technique allows to get functional convergence results for the Kaplan–Meier estimator ([1, 9, 10]; see also

[7, 2]). Note that, the Kaplan–Meier estimator requires the exact event time to be observed, which may fail in practice. In the interval censored data model [27] the event times are not observed exactly. The nonparametric maximum likelihood estimator (NPMLE) for the current status data model can be obtained as a solution of the isotonic regression model [3] using Convex Minorant Algorithm or by using the EM-algorithm [26, 27]. Asymptotic behavior of the NPMLE at any fixed point studied in [11, 16]. Groeneboom & Wellner [16] discussed wide range of asymptotic results on the NPMLE.

The current status right-censored data model discussed in this paper is highly related to the particular case of the current status data with competing risks. The NPMLE and the nonparametric pseudo maximum likelihood estimator (NPPMLE) of parameters from the current status data with competing risks, and the EM-algorithms to get the estimators are given in [17]. Another naïve (ad-hoc) estimator is considered in [18], along with the NPMLE. Consistency and rate of convergence results for the NPMLE are obtained in [14], and weak convergence results are given in [15]. Consistency of the estimators in the current status right-censored data model and the rate of convergence results are obtained in [22].

The current status data and the life table data with a single observation time are particular cases of the model we discuss in this paper. The life table survival data model was widely used at the beginning of survival analysis [4, 6, 8]. The standard life table (actuarial) estimator is generally used to estimate the parameter $F(w_0)$. Breslow & Crowley [5] show that there is no consistent nonparametric estimator of completely unknown distribution function F at the observation time w_0 in the life table survival data model. Nevertheless, in many real cases the asymptotic bias of the standard life table estimator is relatively small [20]. The extended life table estimator that is inconsistent too was investigated in [24].

This work focuses on estimation in current status right-censored data model and investigates properties of nonparametric estimators under small and moderate sample sizes. We consider the NPMLE, the NPPMLE and the naïve estimator, which are obtaining from the corresponding estimators of the baseline current status data model with two competing risks. The maximum likelihood approach and some asymptotic properties of the estimators are discussed in Section 1. The estimation algorithms are displayed in Section 2. Some properties of the estimators obtained by simulations are reported in Section 3, and supplementary tables are postponed to Section 4.

1 The maximum likelihood approach

In this section we display the likelihood function for the interval right-censored data and discuss the nonparametric estimators.

Assume that the failure time T, the censoring time U and the observation time W are independent with the distribution functions F, G and J respectively; $\gamma_T = \sup\{x : F(x) < 1\}$ and $\gamma_G = \sup\{x : G(x) < 1\}$. Let (T_i, U_i, W_i) be a sample from the distribution (T, U, W), and $(W_i, \kappa_i, \kappa_i\delta_i)$ be the observed current status right-censored data, where $X_i = T_i \wedge U_i$, $\delta_i = \mathbb{1}_{\{T_i \le U_i\}}$ and $\kappa_i = \mathbb{1}_{\{X_i \le W_i\}}$, $i = 1, \ldots, n$.

We slightly abuse the notations denoting F, G, J and H for both the distribution functions and the corresponding measures.

The maximum likelihood estimate. Let $\mathcal{Q}$ be the set of nondecreasing nonnegative cadlag functions $Q : \mathbb{R} \to [0, 1]$, such that $\lim_{x\to-\infty} Q(x) = 0$;

$$\mathbb{Q} = \{(Q, Q^*) : Q, Q^* \in \mathcal{Q} \text{ and } Q(x)+Q^*(x) \leq 1, x \in \mathbb{R}\}$$

be the set of parameters of the model. The log-likelihood function for the interval right-censored data is defined for $(Q, Q^*) \in \mathbb{Q}$ as follows:

$$\begin{aligned} LL(Q, Q^*) = &\sum\nolimits_{i=1}^{n} \kappa_i\delta_i \log Q(W_i) \\ &+ \kappa_i(1-\delta_i) \log Q^*(W_i)+(1-\kappa_i) \log(1-Q(W_i)-Q^*(W_i)), \end{aligned} \tag{1}$$

where $Q(x) = \int_0^x (1-G_-)dF = \int_0^x (1-H_-)d\Lambda$, $Q^* \equiv H-Q$ and $H \equiv 1-(1-F)(1-G)$ is the distribution function of the event time X, Λ is the cumulative hazard function corresponding to F restricted to $D_H = \{x : H(x) < 1\}$. A parameter $(\widehat{Q}_n, \widehat{Q}^*_n)$, which maximizes (1) over $(Q, Q^*) \in \mathbb{Q}$ is the NPMLE.

The pseudo maximum likelihood estimate. Let

$$\mathbb{Q}_H = \{(Q, Q^*) \in \mathbb{Q} : Q + Q^* \equiv H\}.$$

The likelihood function (1) can be rewritten as the sum of two terms $LL(W, \kappa, \kappa\delta; F, G) = LL^m(W, \kappa; H) + LL^r(W, \kappa, \kappa\delta; R)$ with

$$LL^m(W, \kappa; H) = \sum\nolimits_{i=1}^{n} \big(\kappa_i \log(H(W_i)) + (1 - \kappa_i) \log(1 - H(W_i))\big)$$

and

$$LL^r(W, \kappa, \kappa\delta; R) = \sum\nolimits_{i=1}^{n} \big(\kappa_i\delta_i \log R(W_i) + \kappa_i(1 - \delta_i) \log(1 - R(W_i))\big),$$

where $R(w) = Q(w)/H(w) = P(\delta = 1|X \leq w) = \int_0^w (1 - H_-)d\Lambda/H(w)$. The functions Q and Q^* can be written as follows:

$$Q(x) = \int_0^x pdH \quad \text{and} \quad Q^*(x) = \int_0^x (1 - p)dH, \tag{2}$$

where $p = \frac{d\Lambda}{d\Lambda^H}$ is the Radon–Nikodym derivative of the measure Λ with respect to Λ^H. Moreover, any measurable function $p : \mathbb{R} \to [0, 1]$ defines the distributions of T and U (possibly improper) under any fixed distribution function H [22]. Let $\widetilde{H}_n$ be the sub distribution function, which maximizes LL^m and $\widetilde{R}_n \equiv \widetilde{Q}_n/\widetilde{H}_n$ maximizes LL^r under $H \equiv \widetilde{H}_n$, and $\widetilde{Q}(x) = \int_0^x pd\widetilde{H}_n$. Then $(\widetilde{Q}, \widetilde{Q}^*)$ such that $\widetilde{Q}^* \equiv \widetilde{H} - \widetilde{Q}$ is the NPPMLE for the parameter (Q, Q^*).

The naïve (ad hoc) approach is based on the separate estimation of the parameters Q and Q^* from the observations with $T < U$ and $T \geq U$ respectively. The naïve estimator $\widehat{\widehat{Q}}_n$ for the parameter Q is obtaining by maximizing

$$\Psi(W, \kappa\delta, Q) = \sum\nolimits_{i=1}^{n} (\kappa_i\delta_i \log Q(W_i) + (1 - \kappa_i\delta_i) \log(1 - Q(W_i))$$

on Q having atoms at the observation points W_i with $\kappa_i\delta_i = 1$, and the naïve estimator $\widehat{\widehat{Q}}_n^*$ for the parameter Q^* is obtaining by maximizing $\Psi(W, \kappa(1-\delta), Q^*)$ on Q^* having atoms at the observation points W_i with $\kappa_i(1-\delta_i)$, $i = 1, \ldots, n$. The naïve estimator can be obtained by the regular convex minorant algorithm from right-censored data analysis. The true disadvantage of the naïve estimator is that the constraint $\widehat{\widehat{Q}}_n + \widehat{\widehat{Q}}_n \le 1$ may fail in the general case.

Recovering the distributions of failure and censoring times. In order to recover the distribution of T from Q and Q^* we use that $\Lambda(x) = \int_0^x (1 - Q_- - Q_-^*)^{-1} dQ$. Hence,

$$S(t) = \prod_{x \le t} \Big(1 - \frac{dQ(x)}{1 - Q(x_-) - Q^*(x_-)}\Big), \tag{3}$$

where $S \equiv 1 - F$. The distribution of censoring time U is determined by the cumulative hazard function $\Lambda^G(x) = \int_0^x \frac{1-F_-}{(1-F)(1-H_-)} dQ^*$ and, therefore, $G(t) = 1 - \prod_{x \le t}\big(1 - d\Lambda^G(x)\big)$. Alternatively, $G(t) = \int_0^t (1-F)^{-1} dQ^*$, $t \in D_H$.

Large sample properties of the estimators. The large sample properties of the nonparametric estimator S_n ($S_n = \widehat{S}_n, \widetilde{S}_n, \widehat{\widehat{S}}_n$) for the distribution of failure time are determined by the large sample properties of the corresponding estimator (Q_n, Q_n^*) for the parameter (Q, Q^*) that is the particular case of the estimator for the current status data with two competing risks model. The uniform consistency and the rate of convergence results for all the estimators S_n were obtained in [22]. In the absolutely continuous case it was proved that under the condition $H \ll J$, for any $\tau < \gamma_F \wedge \gamma_G$

$$\sup_{x \le \tau} |F_n(x) - F(x)| \to 0,$$

as $n \to \infty$ almost sure. The uniform consistency result under the assumption $H \ll J$ remains correct in general case. The condition $H \ll J$ is important, otherwise there is no consistent estimator for the parameter S (see [23, 24]). The rate of convergence in the absolutely continuous case is obtained, under $H \ll J$ and the bounded property $M^{-1} \le \frac{dH}{dJ} \le M$ for some $M > 1$, in the $L_1(J)$ norm restricted to the interval $[0, \tau]$,

$$\|F_n - F\|_{1, J([0,\tau])} = O_P(n^{-1/3} \log^{1/3} n). \tag{4}$$

Remark 1. *(i). The rate of convergence in (4) is obtained from the refined rate of uniform convergence results for the corresponding estimators H_n of the event time distribution function H.*
(ii). We may expect the rate of convergence $O_P(n^{-1/3})$ in (4) taking into account the rate of convergence $O_P(n^{-1/3})$ of the estimators Q_n and Q_n^* to the parameters Q and Q^* in $L_1(J)$ (and even in $L_2(J)$ norm), but the L_p rate of convergence of the estimators Q_n and Q_n^* is not implies the same rate of convergence for the corresponding estimator S_n.
(iii). Local weak convergence theorems for the estimators $(\widehat{Q}_n, \widehat{Q}_n^*)$ and $(\widehat{\widehat{Q}}_n, \widehat{\widehat{Q}}_n^*)$ are given in [15], but there is no way to use these results in order to obtain weak convergence theorem for the corresponding estimators S_n.

2 Estimation algorithms

In this section we discuss algorithms for the NPMLE and the NPPMLE introduced in Section 1. Let $W_{(1)}, \ldots, W_{(r)}$ be the set of observation times in ascending order without replications. The likelihood function (1) can be rewritten in terms of parameters $(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$ with $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_r)$: $\theta_i = Q(W_{(i)})$ and $\boldsymbol{\theta}^* = (\theta_1^*, \ldots, \theta_r^*)$: $\theta_i^* = Q^*(W_{(i)})$, $i = 1, \ldots, r$, as follows:

$$\psi(\boldsymbol{\theta}, \boldsymbol{\theta}^*) = \sum\nolimits_{i=1}^{r} \kappa\delta_{(i)} \log\theta_i + \kappa\bar{\delta}_{(i)} \log\theta_i^* + \bar{\kappa}_{(i)} \log(1 - \theta_i - \theta_i^*),$$

where $\bar{\kappa}_{(i)}$ ($\kappa\delta_{(i)}$, $\kappa\bar{\delta}_{(i)}$) is the total number of observations $W_j = W_{(i)}$ such that $\kappa_j = 0$ ($\kappa_j = 1$ and $\delta_j = 1$, $\kappa_j = 1$ and $\delta_j = 0$, respectively). The optimization problem is to maximize $\psi(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$ on $(\boldsymbol{\theta}, \boldsymbol{\theta}^*) \in \overline{\mathbb{S}}$, where

$$\overline{\mathbb{S}} = \{(\boldsymbol{\theta}, \boldsymbol{\theta}^*) : 0 \le \theta_1 \le \ldots \le \theta_r, 0 \le \theta_1^* \le \ldots \le \theta_r^*, \theta_r + \theta_r^* \le 1\},$$

Let

$$\mathbb{S} = \{(\boldsymbol{\theta}, \boldsymbol{\theta}^*) \in \overline{\mathbb{S}} : \theta_i = \theta_{i-1} \textit{ if } \kappa\delta_{(i)} + \bar{\kappa}_{(i)} = 0 \quad \textit{and} \\ \theta_i^* = \theta_{i-1}^* \textit{ if } \kappa\bar{\delta}_{(i)} + \bar{\kappa}_{(i)} = 0,\ i = 1, \ldots, r\}$$

with the notations $\theta_0 = \theta_0^* = 0$. The NPMLE $(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\theta}}^*)$, which maximizes ψ over $(\boldsymbol{\theta}, \boldsymbol{\theta}^*) \in \mathbb{S}$, is maximizes ψ over $(\boldsymbol{\theta}, \boldsymbol{\theta}^*) \in \overline{\mathbb{S}}$. Moreover, $(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\theta}}^*)$ is uniquely defined, and $\hat{\theta}_r + \hat{\theta}_r^* = 1$ iff $\bar{\kappa}_{(r)} = 0$ [14].

The maximum likelihood estimation requires first to get the NPMLE $(\widehat{Q}, \widehat{Q}^*)$ of the parameter (Q, Q^*) and then recovering the survival function $\widehat{S}_n$ of failure time by formula (3). The first step reduced to the maximum likelihood estimation in the current status data with two competing risks model. The EM-algorithm due to [17] to get the NPMLE for the parameter (Q, Q^*) is working too slow, and one can use the iterated convex minorant (ICM) algorithm (see [12]) based on the characterization of the NPMLE from current status data with competing risk in [14]. Alternatively, the NPMLE for the parameter (Q, Q^*) can be obtained by using the support reduction algorithm [13], which is realized in the R-package *MLEcens* [21].

The pseudo likelihood estimation consists of three steps. At the first step we get the NPMLE $\widetilde{H}_n$ of the parameter H from the interval censored data (X_i, W_i), $i = 1, \ldots, n$. The convex minorant algorithm is a common way to get the maximum likelihood estimator $\widetilde{H}_n$ [16]. At the second step we get the estimator $(\widetilde{Q}_n, \widetilde{Q}_n^*)$, which maximizes LL^r under $\widetilde{Q}_n^* = \widetilde{H}_n - \widetilde{Q}_n$. We study an algorithm to obtain $\widetilde{R}_n \equiv \widetilde{Q}_n / \widetilde{H}_n$ under known $H \equiv \widetilde{H}_n$ from the observed data. Let $W_{(1)}^{**}, \ldots, W_{(m)}^{**}$ be the set of admissible step points of the estimator $\widetilde{R}_n$ in ascending order, including the observation times W_i with $\kappa_i = 1$; $h_1 = H(W_{(1)}^{**}) > 0$ and $h_i = H(W_{(i)}^{**}) - H(W_{(i-1)}^{**}) > 0$ for all $i = 2, \ldots, m$; $\delta_{(i)}^{**} = \sum_{j:W_j = W_{(i)}^{**}} \delta_j$ be the number of observed failures at $W_{(i)}^{**}$, $i = 1, \ldots, m$. It follows from (2) that $\widehat{R}(W_{(s)}^{**}) = \sum_{i=1}^{s} h_i\zeta_i / \sum_{i=1}^{s} h_i$. Then the pseudo-

likelihood function L^r can be rewritten in the following way:

$$L^r(\boldsymbol{\zeta}) = \exp(LL^r(F,G)) = \prod_{s=1}^{m}\Big(\sum_{i=1}^{s} h_i\zeta_i / \sum_{i=1}^{s} h_i\Big)^{\delta^{**}_{(s)}}\Big(\sum_{i=1}^{s} h_i(1-\zeta_i)/\sum_{i=1}^{s} h_i\Big)^{\nu_s-\delta^{**}_{(s)}}$$
$$\cong \prod_{s=1}^{m}\Big(\sum_{i=1}^{s} h_i\zeta_i\Big)^{\delta^{**}_{(s)}}\Big(\sum_{i=1}^{s} h_i(1-\zeta_i)\Big)^{\nu_s-\delta^{**}_{(s)}},$$

where $\boldsymbol{\zeta} = (\zeta_1, \ldots, \zeta_m)$: $\zeta_i = p(W^{**}_{(i)}) \in [0,1]$, ν_i is the total number of observed events at $W^{**}_{(i)}$, $i = 1, \ldots, m$. The estimation problem reduces to maximizing the expression

$$\phi(\boldsymbol{\zeta}) = \sum\nolimits_{s=1}^{m}\Big(\delta^{**}_{(s)}\log\Big(\sum\nolimits_{i=1}^{s} h_i\zeta_i\Big) + (\nu_s - \delta^{**}_{(s)})\log\Big(\sum\nolimits_{i=1}^{s} h_i(1-\zeta_i)\Big)\Big) \qquad (5)$$

over the set of $\boldsymbol{\zeta} \in [0,1]^m$. Finally, at the third step one use the reconstruction formula (3) to obtain the NPPMLE $\widetilde{S}_n$ for S.

3 Simulations

In this section we consider specific designs (DS) to evaluate finite-sample performance of the NPMLE, NPPMLE and the naïve estimator from simulated data. We perform simulations of the current status right-censored data with different rates of observations with known status (failure or censoring) $p_\kappa = P(X \le W)$, which are applicable for estimation of the parameter Q, and different rates of observed failures $p_\delta = P(\delta = 1 | \kappa = 1)$, under the three sample sizes of 200, 500 and 1000. We denote $\Gamma(a,b)$ is the gamma distribution and $\mathcal{W}(a,b)$ is the Weibull distribution with the shape parameter $a > 0$ and scale parameter $b > 0$; $\mathrm{E}(1/b) = \Gamma(1,b)$ is the exponential distribution; $\mathrm{LN}(m,b)$ and $\mathrm{FN}(m,b)$ is the lognormal and the folded-normal distribution with parameters $m \in \mathbb{R}$ and $b > 0$ respectively. The following table 1 collects main features of the experimental designs used for the simulations.

Table 1. Main features of the experimental designs

DS	T	U	W	p_κ	p_δ	DS	T	U	W	p_κ	p_δ
A	$\Gamma(1/2,1)$	$\Gamma(2,1)$	LN(0,1)	0.83	0.91	D	$\Gamma(2,1)$	E(1)	FN(0,1)	0.54	0.19
B	$\Gamma(1/2,1)$	E(1)	FN(0,1)	0.80	0.73	E	$\frac{1}{3}\Gamma(2,\frac{1}{5})+\frac{2}{3}\Gamma(10,\frac{1}{5})$	E(1/2)	E(1)	0.50	0.49
C	$\Gamma(3,1)$	E(1)	E(1)	0.52	0.07	F	$\frac{1}{2}\mathcal{W}(\frac{1}{2},1)+\frac{1}{2}\mathcal{W}(5,1)$	E(1)	E(3/2)	0.48	0.47

The same experimental designs were used in [22] to perform large sample properties of the NPMLE, NPPMLE and its bootstrapped version by simulations.

In order to perform simulations we use R statistical software [25]. The function *computeMLE()* of *MLEcens* package is used to create the MLE $(\widehat{Q}_n, \widehat{Q}^*_n)$ for the parameter (Q,Q). We use the the convex minorant algorithm realization *gcmlcm()* of package *pdrtool* to get the estimator $\widetilde{H}$ for the distribution of the event time H, and the function *lbfgsb3()* of the same name package to solve the optimization problem in (5) under $H \equiv \widetilde{H}$ and obtain the estimator $\widetilde{Q}_n$. Finally, we obtain the

estimators $\widehat{S}_n$, $\widetilde{S}_n$ and $\widetilde{\widetilde{S}}_n$ for the survival function S of failure time from $(\widehat{Q}_n, \widehat{Q}^*_n)$, $(\widetilde{Q}_n, \widetilde{Q}^*_n)$ by the reconstruction formula (3).

For the NPMLE, NPPMLE and naïve estimators we display the estimation bias (Section 4, Table I) and the mean absolute estimation error (Section 4, Table II) at the quartiles (Q25,MED,Q75) and 95%-quantile Q95 of the failure time distribution, as well as the supremum $\sup_{x\in[0,Q]} |S_n(x) - S(x)|$ (Section 4, Table III) and the L_1 norm adjusted to the interval length $\|S_n - S\|_{1,[0,Q]}/J((0,Q])$ (Section 4, Table IV) restricted to the interval $[0, Q]$ for Q =Q25,MED,Q75,Q95. The results are obtained separately by using 10^4 replications.

First, we note that the finite sample performance of the estimators is highly related to the experimental design features. The designs A and B display very good approximation quality for MED$-Q95$ quantiles, but there is an obvious problems in the estimation of the survival distribution of failure time at first quartile $Q25$, especially under the experimental design A because of $\frac{dQ}{dJ}(w) \to \infty$ as $w \to 0_+$. On the other hand, the number of observations is insufficient to get good enough nonparametric estimates under the designs C and D having a very small rate of observed failures. All the estimators display good enough finite sample performance under the designs E and F with the bimodal distributions of failure time. The $L_1(J)$ divergence display quite small estimation error for all the designs except the design A, and the uniform norm divergence is too high under these sample sizes. Moreover, both the $L_1(J)$ and the uniform estimation errors are not highly dependent of the population sizes from 200 to 1000.

Roughly, the nonparametric estimators show very similar finite sample performance for each of the designs. More careful look at the results allows us to give some preference to the NPMLE, which displays a little bit smaller divergence in almost all the cases. In most of cases the NPPMLE performs a little bit better results then the naïve estimator, but it displays a huge bias (overestimation of the survival function) at Q75 and Q95 quantile points that should be explained by accumulation of the bias and the estimation error appears under estimation of the event time distribution H and the competing risks components (Q, Q^*) under fixed $H \equiv \widetilde{H}$ in the adverse experimental conditions.

4 Supplementary tables

Table I. The estimation bias

		NPMLE				NPPMLE				Naïve			
DS	N	Q25	MED	Q75	Q95	Q25	MED	Q75	Q95	Q25	MED	Q75	Q95
A	200	0.2314	0.0309	6.9E-4	-0.009	0.2314	0.0282	0.0013	-0.0138	0.2337	0.0885	0.0947	0.137
	500	0.2088	0.0141	9.4E-5	-0.0128	0.2086	0.012	0.0012	-0.007	0.2141	0.0768	0.1002	0.1505
	1000	0.1804	0.0085	1.5E-4	-0.0101	0.1803	0.0066	7.9E-4	-1.1E-4	0.1899	0.0735	0.1029	0.1565
B	200	0.0716	0.0177	-0.0045	-0.0029	0.0602	0.0171	0.0150	0.0132	0.1048	0.1112	0.1711	0.1664
	500	0.0314	0.0087	-0.0012	-0.0139	0.0242	0.0099	0.0112	0.0022	0.073	0.105	0.1721	0.2128
	1000	0.0192	0.0052	-0.0010	-0.0184	0.0151	0.0074	0.0085	-0.0025	0.0624	0.1019	0.171	0.2353
C	200	-0.0439	-0.0914	0.0561	0.2349	0.0112	0.1732	0.4094	0.6086	0.0475	0.0558	0.0447	0.1064
	500	-0.0277	-0.1052	0.0146	0.1959	0.0018	0.0932	0.3181	0.5163	0.0735	0.1515	0.1711	0.0989
	1000	-0.0173	-0.0926	-0.0097	0.173	0.0052	0.0429	0.2500	0.4471	0.0781	0.1853	0.2528	0.1257
D	200	-0.0081	-0.0932	-0.0318	0.1589	0.0106	0.0238	0.1893	0.3865	0.0554	0.073	-0.0657	0.0987
	500	-0.0051	-0.0605	-0.0669	0.1193	0.0084	0.0027	0.1278	0.3222	0.0616	0.1278	0.0042	0.0882
	1000	-0.0045	-0.0337	-0.0858	0.0945	0.0062	0.0015	0.0852	0.2776	0.0627	0.1492	0.0865	0.0802
E	200	0.0146	-0.0392	-0.0700	-0.0097	0.0176	-0.0045	0.0594	0.1742	0.0608	0.1036	0.2079	0.2467
	500	0.0084	-0.0183	-0.0510	-0.022	0.0106	-0.0010	0.0383	0.1169	0.059	0.1200	0.2311	0.3046
	1000	0.0057	-0.0102	-0.0290	-0.0248	0.0086	-3.1E-4	0.0237	0.0796	0.0584	0.1265	0.2359	0.3278
F	200	-0.0075	-0.0382	-0.0378	0.053	-0.0067	0.0057	0.0851	0.1612	0.0449	0.1289	0.2572	0.0964
	500	-0.0047	-0.021	-0.0269	0.043	-0.0039	7.5E-4	0.0491	0.0852	0.0486	0.135	0.2653	0.0951
	1000	-0.0032	-0.0121	-0.0166	0.0347	-0.0020	0.0018	0.0306	0.0526	0.0508	0.1377	0.265	0.0946

Table II. The absolute error

		NPMLE				NPPMLE				Naïve			
DS	N	Q25	MED	Q75	Q95	Q25	MED	Q75	Q95	Q25	MED	Q75	Q95
A	200	0.2413	0.0946	0.0560	0.0355	0.2412	0.0945	0.0583	0.0537	0.2404	0.1161	0.1096	0.1446
	500	0.2257	0.0637	0.0402	0.0288	0.2256	0.0643	0.0425	0.0424	0.2244	0.0934	0.1059	0.1519
	1000	0.2024	0.0485	0.032	0.0236	0.2019	0.0492	0.034	0.0328	0.2016	0.083	0.1055	0.1569
B	200	0.1251	0.0774	0.0649	0.0487	0.1215	0.0796	0.074	0.0804	0.1278	0.1214	0.1749	0.1887
	500	0.0770	0.0561	0.0451	0.0372	0.0761	0.0591	0.0521	0.0645	0.0882	0.1099	0.173	0.2198
	1000	0.0571	0.0442	0.0350	0.0327	0.0573	0.0475	0.0416	0.0540	0.0726	0.1041	0.1713	0.2375
C	200	0.1534	0.2540	0.2696	0.2803	0.1341	0.2809	0.4635	0.6194	0.1424	0.2826	0.3063	0.1780
	500	0.1009	0.2052	0.2188	0.2390	0.1023	0.2165	0.3763	0.5278	0.1065	0.2352	0.3216	0.1689
	1000	0.0741	0.1687	0.1889	0.2143	0.0804	0.1747	0.3103	0.4584	0.0969	0.2214	0.3367	0.1899
D	200	0.0784	0.1706	0.1811	0.2009	0.0801	0.179	0.3076	0.4101	0.0894	0.1893	0.2463	0.1647
	500	0.0561	0.1205	0.1520	0.1619	0.0601	0.1317	0.2474	0.3459	0.0770	0.1694	0.2437	0.1544
	1000	0.0442	0.0855	0.1406	0.1384	0.0485	0.1045	0.2086	0.3014	0.0719	0.1671	0.2432	0.1467
E	200	0.0574	0.1186	0.1606	0.0643	0.0590	0.1116	0.1979	0.2216	0.0730	0.1327	0.2306	0.2634
	500	0.0404	0.0830	0.1172	0.0522	0.0425	0.0830	0.1334	0.1636	0.0639	0.1297	0.236	0.3083
	1000	0.0321	0.0653	0.0867	0.0464	0.034	0.0678	0.0979	0.1229	0.0609	0.1302	0.2376	0.3287
F	200	0.0558	0.1321	0.1258	0.0919	0.0559	0.1256	0.1851	0.2089	0.0638	0.1509	0.2662	0.1498
	500	0.0395	0.0935	0.0898	0.0719	0.0412	0.0940	0.1186	0.1336	0.0567	0.1434	0.2672	0.1481
	1000	0.0312	0.0715	0.0693	0.0608	0.0328	0.0763	0.0871	0.0997	0.0547	0.1417	0.2654	0.1472

Table III. The supremum norm divergence

		NPMLE				NPPMLE				Naïve			
DS	N	Q25	MED	Q75	Q95	Q25	MED	Q75	Q95	Q25	MED	Q75	Q95
A	200	0.2490	0.3506	0.3512	0.3512	0.2490	0.3505	0.351	0.3511	0.2487	0.3515	0.3537	0.3564
	500	0.2466	0.3059	0.3060	0.3060	0.2466	0.3058	0.3059	0.3059	0.2462	0.3069	0.3084	0.3131
	1000	0.2423	0.2767	0.2767	0.2767	0.2424	0.2767	0.2767	0.2767	0.2420	0.2779	0.2795	0.2869
B	200	0.2036	0.2327	0.2381	0.2393	0.2025	0.2305	0.2373	0.2438	0.1998	0.2429	0.2752	0.3167
	500	0.1661	0.1805	0.1836	0.1841	0.1645	0.1801	0.1846	0.1888	0.1650	0.1988	0.2380	0.2940
	1000	0.1379	0.1480	0.1499	0.1502	0.137	0.1486	0.1520	0.1549	0.1401	0.1734	0.2182	0.2849
C	200	0.2126	0.3879	0.4621	0.5364	0.1836	0.3432	0.5240	0.6872	0.183	0.3607	0.4755	0.5256
	500	0.1524	0.323	0.3913	0.457	0.1475	0.2897	0.4451	0.5963	0.1394	0.2901	0.4287	0.4942
	1000	0.1205	0.2683	0.3449	0.4044	0.1230	0.2504	0.3845	0.5249	0.1252	0.2655	0.4136	0.4951
D	200	0.1323	0.2673	0.3465	0.3960	0.1297	0.2530	0.3916	0.5178	0.1306	0.2573	0.3777	0.4199
	500	0.1014	0.1957	0.2978	0.3351	0.1038	0.1995	0.3258	0.4421	0.1101	0.2207	0.3453	0.3898
	1000	0.0837	0.1513	0.2663	0.2972	0.0877	0.1662	0.2832	0.3909	0.0998	0.2084	0.3335	0.3794
E	200	0.1189	0.1953	0.2862	0.3005	0.1171	0.1856	0.2860	0.3637	0.1227	0.1908	0.3013	0.3913
	500	0.0909	0.1447	0.2193	0.2340	0.0919	0.1447	0.2189	0.2800	0.1021	0.1732	0.2893	0.3871
	1000	0.0748	0.1184	0.1751	0.1905	0.0768	0.1217	0.1803	0.2255	0.0913	0.1663	0.2819	0.383
F	200	0.1146	0.2122	0.2812	0.2905	0.1132	0.2002	0.2895	0.3578	0.1161	0.2022	0.3288	0.4004
	500	0.0878	0.1623	0.2149	0.2208	0.0880	0.1616	0.2223	0.2632	0.0964	0.1859	0.3120	0.3855
	1000	0.072	0.1324	0.1730	0.1776	0.0726	0.1365	0.1852	0.2120	0.0859	0.1787	0.3019	0.3768

Table IV. The L_1 norm divergence

		NPMLE				NPPMLE				Naïve			
DS	N	Q25	MED	Q75	Q95	Q25	MED	Q75	Q95	Q25	MED	Q75	Q95
A	200	0.2159	0.1525	0.0854	0.0653	0.2159	0.1541	0.0876	0.0709	0.2154	0.1643	0.1188	0.1224
	500	0.2096	0.1001	0.0590	0.0459	0.2095	0.1015	0.0610	0.0493	0.2085	0.1195	0.1018	0.1152
	1000	0.1980	0.0735	0.0452	0.0353	0.1979	0.0747	0.0469	0.0379	0.1971	0.0972	0.0944	0.1130
B	200	0.1207	0.0979	0.0803	0.0717	0.1194	0.0979	0.0831	0.0849	0.1184	0.1174	0.1366	0.1675
	500	0.0876	0.0688	0.0568	0.0510	0.0864	0.0699	0.0604	0.0607	0.0894	0.0951	0.1246	0.1653
	1000	0.0663	0.0528	0.0441	0.0395	0.0656	0.0544	0.0478	0.0476	0.0718	0.0845	0.119	0.1642
C	200	0.0337	0.0541	0.0645	0.0683	0.0307	0.0496	0.0651	0.0734	0.031	0.0512	0.0642	0.0677
	500	0.0231	0.0381	0.0469	0.0501	0.0226	0.0375	0.0498	0.0568	0.0231	0.0392	0.0515	0.0556
	1000	0.0177	0.0290	0.0367	0.0396	0.0181	0.0302	0.0401	0.0461	0.0199	0.0348	0.0470	0.0518
D	200	0.0391	0.0596	0.0704	0.0712	0.0377	0.0592	0.0741	0.0759	0.0400	0.0643	0.0787	0.0797
	500	0.0281	0.0414	0.0508	0.0515	0.0283	0.0438	0.0552	0.0567	0.0318	0.0541	0.0673	0.0684
	1000	0.0220	0.0318	0.0396	0.0403	0.0227	0.0349	0.0442	0.0455	0.0277	0.0500	0.0633	0.0644
E	200	0.0437	0.0572	0.0667	0.0703	0.0426	0.0566	0.0665	0.0771	0.0468	0.0659	0.0778	0.0905
	500	0.0313	0.0405	0.0470	0.0504	0.0314	0.0413	0.0480	0.0556	0.0375	0.058	0.0707	0.0854
	1000	0.0246	0.0317	0.0365	0.0394	0.0250	0.0329	0.0381	0.0434	0.0326	0.0547	0.0678	0.0833
F	200	0.0499	0.0610	0.0697	0.0741	0.0491	0.0597	0.0699	0.0946	0.0562	0.0677	0.0826	0.1133
	500	0.0348	0.0432	0.0493	0.0524	0.0350	0.0439	0.0506	0.0652	0.046	0.0585	0.0739	0.1096
	1000	0.0270	0.0338	0.0383	0.0408	0.0275	0.0349	0.0401	0.0500	0.0414	0.0549	0.0704	0.1089

Acknowledgements

This research was supported in part by St.-Petersburg State University (grants No 1.52.1647.2016).

References

[1] Aalen, O. (1978). Nonparametric inference for a family of counting processes. *The Annals of Statistics* **6**(4), pp. 701–726.

[2] Andersen, P. K., Borgan, Ø., R. D. Gill and Keiding, N. (1993). *Statistical Models Based on Counting Processes.* New York: Springer-Verlag.

[3] Ayer, M., Brunk, H. D., Ewing, G. M., Reid, W. T. and Silverman, E. (1955). An empirical distribution function for sampling with incomplete information. *The Annals of Mathematical Statistics* **26**(4), 641–647.

[4] Berkson and Gage (1950). Calculation of survival rates for cancer. *Proceedings of Staff Meetings of the Mayo Clinic*, Vol. **25**, pp. 270–286.

[5] Breslow, N. and Crowley, J. (1974) A Large Sample Study of the Life Table and Product Limit Estimates Under Random Censorship. *The Annals of Statistics*, Vol. **2**(3), pp. 437–453.

[6] Cutler and Ederer (1958). Maximum utilization of the life table method in analyzing survival. *J. Chron.Dis.*, . Vol. **8**, pp. 699–712.

[7] Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis.* New York: John Wiley & Sons.

[8] Gehan (1969). Estimating survival function from the life table. *J.Chron.Dis.*, Vol. **21**, pp. 629–644.

[9] Gill, R. D. (1980a). *Censoring and stochastic integrals.* Mathematical Centre Tracts 124. Amsterdam: Mathematisch Centrum. V, 178 p.

[10] Gill, R. D. (1980b). Nonparametric estimation based on censored observations of a markov renewal process. *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete* **53**(1), 97–116.

[11] Groeneboom, P. (1991). *Nonparametric maximum likelihood estimators for interval censoring and deconvolution.* Technical Report 378, Department of Statistics, Stanford University.

[12] Groeneboom, P. and Jongbloed, G. (2014). *Nonparametric Estimation under Shape Constraints: Estimators, Algorithms and Asymptotics.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

[13] Groeneboom, P., G. Jongbloed, and J. Wellner (2008). The support reduction algorithm for computing non-parametric function estimates in mixture models. *Scandinavian Journal of Statistics* **35**(3), 385–399.

[14] Groeneboom, P., M. Maathuis, and J. Wellner (2008a). Current status data with competing risks: Consistency and rates of convergence of the mle. *Ann. Statist.* **36**(3), 1031–1063.

[15] Groeneboom, P., Maathuis, M. and Wellner, J. (2008b). Current status data with competing risks: Limiting distribution of the mle. *Ann. Statist.* **36**(3), 1064–1089.

[16] Groeneboom, P. and J. A. Wellner (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation.* DMV Seminar Band 19. Basel: Birkhauser Verlag.

[17] Hudgens, M. G., Satten, G. and Longini, I. (2001). Nonparametric maximum likelihood estimation for competing risks survival data subject to interval censoring and truncation. *Biometrics* **57**(1), 74–80.

[18] Jewell, N., van der Laan, M. and Henneman, T. (2003). Nonparametric estimation from current status data with competing risks. *Biometrika* **90**(1), 183–197.

[19] Kaplan, E. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* **53**(282), 457–481.

[20] Lawless, J. F. (2002). *Statistical Models and Methods for Lifetime Data, 2-nd edition.* Wiley Series in Probability and Statistics. John Wiley & Sons.

[21] Maathuis, M. (2013). *MLEcens: Computation of the MLE for bivariate (interval) censored data.* R package version 0.1-4.

[22] Malov S.V. (2019). Nonparametric estimation for current status interval right-censored data model. *Statistica Neerlandica*, 1–21. Published online https://doi.org/10.1111/stan.12180.

[23] Malov S.V. & O'Brien S.J. (2015). On Survival Categorical Methods Based on an Extended Actuarial Estimator. In *Applied Methods of Statistical Analysis. Nonparametric Approach.* Proceedings of the International Workshop AMSA'15 (Novosibirsk & Belokurikha, September 14-19, 2015), pp. 147–156.

[24] Malov S.V. & O'Brien S.J. (2018). Life Table Estimator Revisited. *Communication in Statistics – Theory and Methods* **47**(9), 2126–2133.

[25] R Core Team (2014). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

[26] Turnbull, B. W. (1974). Nonparametric estimation of a survivorship function with doubly censored data. *Journal of the American Statistical Association* **69**, 169–173.

[27] Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society.* Series B **38**, 290–295.

Survival function estimation from fixed design regression model in the presence of dependent random censoring

Abdurakhim A. Abdushukurov
Branch of Moscow State University in Tashkent, Uzbekistan
e-mail: a_abdushukurov@rambler.ru

Abstract

The aim of paper is considering the problem of estimation of conditional survival function in the case of right random censoring with presence of covariate. For proposed estimator we prove an almost sure representation result with rate and weak convergence results to the Gaussian processes.

Keywords: Survival function, random censoring, covariate, Gaussian processes.

Introduction

The aim of paper is considering the problem of estimation of conditional survival function in the case of right random censoring with presence of covariate. Let's consider the case when the support of covariate C is the interval $[0,1]$ and we describe our results on fixed design points $0 \leq x_1 \leq x_2 \leq ... \leq x_n \leq 1$ at which we consider responses (survival or failure times) $X_1, ..., X_n$ and censoring times $Y_1, ..., Y_n$ of identical objects, which are under study. These responses are independent and nonnegative random variables (r.v.-s) with conditional distribution function (d.f.) at x_i,$F_{x_i}(t) = P(X_i \leq t/C_i = x_i)$. They are subjected to random right censoring, that is for X_i there is a censoring variable Y_i with conditional d.f. $G_{x_i}(t) = P(Y_i \leq t/C_i = x_i)$ and at $n-$th stage of experiment the observed data is

$$S^{(n)} = \{(Z_i, \delta_i, C_i), 1 \leq i \leq n\},$$

where $Z_i = min(X_i, Y_i), \delta_i = I(X_i \leq Y_i)$ with $I(A)$ denoting the indicator of event A. Note that in sample S^n r.v. X_i is observed only when $\delta_i = 1$. Commonly, in survival analysis to assume independence between the r.v.-s X_i and Y_i conditional on the covariate C_i. But, in some practical situations, this assumption does not hold. Therefore, in this article we consider a dependence model in which dependence structure is described through copula function.

1 Estimation of survival function

Let

$$S_x(t_1, t_2) = P(X_x > t_1, Y_x > t_2),\ t_1, t_2 \geq 0,$$

the joint survival function of the response X_x and the censoring variable Y_x at x. Then the marginal survival functions are ${S_x}^X(t) = 1 - F_x(t) = S_x(t,0)$ and ${S_x}^Y(t) = 1 - G_x(t) = S_x(0,t), t \leq 0$. We suppose that the marginal d.f.-s F_x and G_x are continuous. Then according to the Theorem of Sclar (see, [1]), the joint survival function $S_x(t_1,t_2)$ can be expressed as

$$S_x(t_1,t_2) = C_x(S_x^X(t_1), S_x^Y(t_2)), \ t_1,t_2 \geq 0, \tag{1}$$

where $C_x(u,v)$ is a known copula function depending on x, S_x^X and S_x^Y in a general way. We consider estimator of d.f. F_x which is equivalent to the relative-risk power estimator [2,3] under independent censoring case.

Assume that at the fixed design value $x \in (0,1), C_x$ in (1) is Archimedean copula, i.e.

$$S_x(t_1,t_2) = \varphi_x^{[-1]}(\varphi_x(S_x^X(t_1)) + \varphi_x(S_x^Y(t_2))), \ t_1,t_2 \geq 0, \tag{2}$$

where, for each $x, \varphi_x : [0,1] \to [0,+\infty]$ is a known continuous, convex, strictly decreasing function with $\varphi_x = 0$. We assume that copula generator function φ_x is strict, i.e. $\varphi_x(0) = \infty$ and φ_x^{-1} is a inverse of φ_x. From (2), it follows that

$$P(Z_x > t) = 1 - H_x(t) = \overline{H_x(t)} = {S_x}^Z(t) = S_x(t,t) =$$

$$= \varphi_x^{-1}(\varphi_x(S_x^X(t)) + \varphi_x(S_x^Y(t))), \ t \geq 0, \tag{3}$$

Let $H_x^{(1)}(t) = P(Z_x \leq t, \delta_x = 1)$ be a subdistribution function and $\Lambda_x(t)$ is crude hazard function of r.v. X_x subjecting to censoring by Y_x,

$$\Lambda_x(dt) = \frac{P(X_x \in dt, X_x \leq Y_x)}{P(X_x \geq t, Y_x \geq t)} = \frac{H_x^{(1)}(dt)}{S_x^Z(t-)}. \tag{4}$$

From (4) one can obtain following expression of survival function S_x^X :

$$S_x^X(t) = \varphi_x^{-1}[-\int_0^t \varphi_x^{'}(S_x^Z(u))dH_x^{(1)}(u)], \ t \geq 0. \tag{5}$$

In order to constructing the estimator of S_x^X according to representation (5), we introduce smoothed estimators of $S_x^Z, H_x^{(1)}$ and regularity conditions for them. We We use the Gasser-Müller weights

$$w_{ni}(x,h_n) = \frac{1}{q_n(x,h_n)} \int_{x_{i-1}}^{x_i} \frac{1}{h_n}\pi(\frac{x-z}{h_n})dz, \ i = 1,...,n, \tag{6}$$

with

$$q_n(x,h_n) = \int_0^{x_n} \frac{1}{h_n}\pi(\frac{x-z}{h_n})dz,$$

where $x_0 = 0$, π is a known probability density function (kernel) and $\{h_n, n \geq 1\}$ is a sequence of positive constants, tending to zero as $n \to \infty$, called bandwidth sequence. Let's introduce the weighted estimators of H_x, S_x^Z and $H_x^{(1)}$ respectively as

$$H_{xh}(t) = \sum_{i=1}^{n} w_{ni}(x, h_n) I(Z_i \leq t),\ S_{xh}^Z(t) = 1 - H_{xh}(t), H_{xh}^{(1)}(t) =$$

$$= \sum_{i=1}^{n} w_{ni}(x, h_n) I(Z_i \leq t, \delta_i = 1). \tag{7}$$

Then pluggin in (5) estimators (6) and (7) we obtained the following intermediate estimator of S_x^X :

$$S_{xh}^X(t) = 1 - F_{xh}(t) = \varphi_x^{-1}[-\int_0^t \varphi_x^{'}(S_x^Z(u)) dH_x^{(1)}(u)],\ t \geq 0.$$

In this work we propose the next extended analogue of estimator introduced in [2,3]:

$$\widehat{S}_{xh}^X(t) = \varphi_x^{-1}[\varphi(S_{xh}^Z(t)) \cdot \mu_{xh}(t)] = 1 - \widehat{F}_{xh}(t), \tag{8}$$

where $\mu_{xh}(t) = \varphi(S_{xh}^X(t))/\varphi(\widetilde{S}_{xh}^Z(t)),\ \ \varphi(S_{xh}^X(t)) = -\int_0^t \varphi_x^{'}(S_{xh}^Z(u)) dH_{xh}^{(1)}(u),$

$\varphi(\widetilde{S}_{xh}^Z(t)) = -\int_0^t \varphi_x^{'}(S_{xh}^Z(u)) dH_{xh}(u)$. In order to investigate the estimate (8) we introduce some conditions. For the design points $x_1, ..., x_n$, denote $\underline{\Delta_n} = \min_{1 \leq i \leq n} (x_i - x_{i-1}),\ \overline{\Delta_n} = \max_{1 \leq i \leq n} (x_i - x_{i-1})$.

For the kernel π, let $\|\pi\|_2^2 = \int_{-\infty}^{\infty} \pi^2(u) du,\ m_\nu(\pi) = \int_{-\infty}^{\infty} u^\nu \pi(u) du,\ \nu = 1, 2.$

Moreover, we use next assumptions on the design and on the kernel function:
(A1) As $n \to \infty, x_n \to 1, \underline{\Delta_n} = O(\frac{1}{n}), \overline{\Delta_n} - \underline{\Delta_n} = o(\frac{1}{n})$.
(A2) π is a probability density function with compact support $[-M, M]$ for some $M > 0$, with $m_1(\pi) = 0$ and $|\pi(u) - \pi(u^{'})| \leq C(\pi)|u - u^{'}|$, where $C(\pi)$ is some constant.

Let $T_{H_x} = inf\{t \geq 0 : H_x(t) = 1\}$. Then $T_{H_x} = min(T_{F_x}, T_{G_x})$. For our results we need some smoothnees conditions on functions $H_x(t)$ and $H_x^{(1)}(t)$. We formulate them for a general (sub)distribution function $N_x(t), 0 \leq x \leq 1, t \in R$ and for a fixed $T > 0$.
(A3) $\frac{\partial^2}{\partial x^2} N_x(t) = \ddot{N}_x(t)$ exists and is continuous in $(x, t) \in [0, 1] \times [0, T]$.
(A4) $\frac{\partial^2}{\partial t^2} N_x(t) = N_x^{''}(t)$ exists and is continuous in $(x, t) \in [0, 1] \times [0, T]$.
(A5) $\frac{\partial^2}{\partial x \partial t} N_x(t) = \dot{N}^{'}{}_x(t)$ exists and is continuous in $(x, t) \in [0, 1] \times [0, T]$.
(A6) $\frac{\partial \varphi_x(u)}{\partial u} = \varphi_x^{'}(u)$ and $\frac{\partial^2 \varphi_x(u)}{\partial u^2} = \varphi_x^{''}(u)$ are Lipschitz in the $x-$direction with a

bounded Lipschitz constant and $\frac{\partial^3 \varphi_x(u)}{\partial u^3} = \varphi_x^{'''}(u)$ exists and is continuous in $(x,u) \in [0,1] \times (0,1]$.

2 Asymptotic properties of estimators

We derive an almost sure representation result with rate.

Theorem 1. Assume (A1), (A2), $H_x(t)$ and $H_x^{(1)}(t)$ satisfy (A3)-(A5) in $[0,T]$ with $T < T_{H_x}$, φ_x satisfies (A6) and $h_n \to 0$, $\frac{logn}{nh_n} \to 0$, $\frac{nh_n^5}{logn} = O(1)$. Then, as $n \to \infty$,

$$\widehat{F}_{xh}(t) - F_x(t) = \sum_{i=1}^{n} w_{ni}(x, h_n)\Psi_{tx}(Z_i, \delta_i) + r_n(t),$$

where

$$\Psi_{tx}(Z_i,\delta_i) = \frac{-1}{\varphi_x^{'}(S_x^X(t))}[\int_0^t \varphi_x^{''}(S_x^Z(u))(I(Z_i \le u) - H_x(u))dH_x^{(1)}(u) -$$

$$-\varphi_x^{'}(S_x^Z(t))(I(Z_i \le t, \delta_i = 1) - H_x^{(1)}(t)) -$$

$$-\int_0^t \varphi_x^{''}(S_x^Z(u))(I(Z_i \le u, \delta_i = 1) - H_x^{(1)}(u))dH_x(u)],$$

and

$$\sup_{0 \le t \le T} |r_n(t)| \overset{a.s.}{=} O((\frac{\log n}{nh_n})^{3/4}).$$

The weak convergence of the empirical process $(nh_n)^{1/2}\{\widehat{F}_{xh}(\cdot) - F_x(\cdot)\}$ in the space $l^\infty[0,T]$ of uniformly bounded functions on $[0,T]$, endowed with the uniform topology is the contents of the next theorem.

Theorem 2. Assume (A1), (A2), $H_x(t)$ and $H_x^{(1)}(t)$ satisfy (A3)-(A5) in $[0,T]$ with $T < T_{H_x}$, and that φ_x satisfies (A6).
(I) If $nh_n^5 \to 0$ and $\frac{(logn)^3}{nh_n} \to 0$, then, as $n \to \infty$,

$$(nh_n)^{1/2}\{\widehat{F}_{xh}(\cdot) - F_x(\cdot)\} \Rightarrow \mathbf{W}_x(\cdot) \; in \; l^\infty[0,T].$$

(II)If $h_n = Cn^{-1/5}$ for some $C > 0$, then, as $n \to \infty$,

$$(nh_n)^{1/2}\{\widehat{F}_{xh}(\cdot) - F_x(\cdot)\} \Rightarrow \mathbf{W}_x^*(\cdot) \; in \; l^\infty[0,T],$$

where $\mathbf{W}_x(\cdot)$ and $\mathbf{W}_x^*(\cdot)$ are Gaussian processes with means

$$E\mathbf{W}_x(t) = 0, E\mathbf{W}_x^*(t) = a_x(t),$$

and same covariance

$$Cov(\mathbf{W}_x(t), \mathbf{W}_x^(s)) = Cov(\mathbf{W}_x^*(t), \mathbf{W}_x^*(s)) = \Gamma_x(t,s),$$

with

$$a_x(t) = \frac{-C^{5/2} m_2(\pi)}{2\varphi_x^{'}(S_x^X(t))} \int_0^t [\varphi_x^{''}(S_x^Z(u)) \ddot{H}_x(u) dH_x^{(1)}(u) - \varphi_x^{'}(S_x^Z(u)) d\ddot{H}_x^{(1)}(u)\,],$$

and

$$\Gamma_x(t,s) = \frac{\|\pi\|_2^2}{\varphi_x^{'}(S_x^X(t))\varphi_x^{'}(S_x^X(s))} \{ \int_0^{\min(t,s)} (\varphi_x^{'}(S_x^Z(z)))^2 dH_x^{(1)}(z) +$$

$$+ \int_0^{\min(t,s)} [\varphi_x^{''}(S_x^Z(w)) S_x^Z(w) + \varphi_x^{'}(S_x^Z(w))] \int_0^w \varphi_x^{''}(S_x^Z(y)) dH_x^{(1)}(y) dH_x^{(1)}(w) +$$

$$+ \int_0^{\min(t,s)} \varphi_x^{''}(S_x^Z(w)) \int_w^{\max(t,s)} (\varphi_x^{''}(S_x^Z(y)) S_x^Z(y) + \varphi_x^{'}(S_x^Z(y))) dH_x^{(1)}(y) dH_x^{(1)}(w) -$$

$$- \int_0^t [\varphi_x^{''}(S_x^Z(y)) S_x^Z(y) + \varphi_x^{'}(S_x^Z(y))] dH_x^{(1)}(y) \cdot$$

$$\cdot \int_0^s [\varphi_x^{''}(S_x^Z(w)) S_x^Z(w) + \varphi_x^{'}(S_x^Z(w))] dH_x^{(1)}(w) \}.$$

References

[1] Nelsen R.B. (1999). *An Introduction to Copulas*. Springer, New York.

[2] Abdushukurov A.A. (1998). Nonparametric estimation of distribution function based on relative risk function. *Commun. Statist.: Theory and Methods.* Vol. **27**, No.**8**, pp. 1991-2012.

[3] Abdushukurov A.A. (1999). On nonparametric estimation of reliability indices by censored samples. *Theory Probab. Appl.* Vol. bf 43, No. 1, pp. 3-11.

Asymptotics of chi-square test based on the likelihood ratio statistics under random censoring from both sides

Nargiza S. Nurmukhamedova
National University of Uzbekistan, Tashkent, Uzbekistan
e-mail: rasulova_nargiza@mail.ru

Abstract

We consider research two statistics for testing the composite hypotheses in a model of random censorship from both sides, which have in limit a chi-square distribution with appropriate degrees of freedom. First one is the generalized chi-square statistics, for the construction of which we use the power estimate distributions of function. The second statistics is twice the logarithm of the likelihood ratio statistics of model of random censorship from both sides. Both of these statistics can be used to construct an asymptotic tests of chi-square type for the composite hypotheses.

Keywords: Chi-square statistics, likelihood ratio statistics ,maximum likelihood estimate, random censoring.

Introduction

Chi-square statistics occupies an important place in hypothesis testing theory. Having more than a century of history, Chi-square statistics to date has various modifications and generalizations. Among the vast amount of literature should be allocated monograph authors [5], which describes the theories, methods and applications of various statistics such as chi-square to build the criteria of consent. It is known that these statistics use jumps of the empirical distribution function in their structure, i.e. relative frequencies as an estimate for the probability of observations falling into the grouping intervals. In the case of incomplete observations, the empirical distribution function is an untenable estimate for the unknown distribution function (d.f) and therefore, instead of the empirical distribution function, different estimates have to be used for the d.f. whose structures depend on the models under consideration (see [1] for details). Thus, in the work [3] with random censorship on the right, the Chi-square statistics of the consent criterion is constructed and studied using a nonparametric Kaplan-Meier estimate [4] for the d.f.. In [2] established properties of locally asymptotically normality for likelihood ratio statistics (LRS) in competing risks model under random censoring. In this paper, in a general model of incomplete data constructed and investigated chi-square type statistics. Proved a result that can be used to construct the chi-square test based on the asymptotics of the doubled logarithm of the LRS.

Let $\{(X_i, Y_{1i}, Y_{2i}), i \geq 1\}$ sequence of independent and identically distributed (i.i.d) random vectors with mutually independent components and marginal d.f.-s F and G_k for random variables (r.v.) X_i and Y_{ki}, $k = 1, 2$; $i \geq 1$, respectively. Consider the case, when r.v. X_i subject to random censoring from both

sides by variables Y_{ki} . On n- th stage of the experiment we observe the sample of size n: $S^{(n)} = \{(Z_i, \delta_{0i}, \delta_{1i}, \delta_{2i}), 1 \le i \le n\}$, where $Z_i = Y_{1i} \vee (X_i \wedge Y_{2i})$ $\delta_{0i} = I(X_i \wedge Y_{2i} < Y_{1i})$, $\delta_{1i} = I(Y_{1i} \le X_i \le Y_{2i})$, $\delta_{2i} = I(Y_{1i} \le Y_{2i} < X_i)$. Here for numbers a and b: $a \wedge b = \min(a, b)$, $a \vee b = \max(a, b)$. In a sample $S^{(n)}$ r.v. X_i observed only when$\delta_{1i} = 1$. In this model of random censorship from the both sides of the problem consists in estimating of conditional survival function $1 - F_\tau(x) = P(X_i \ge x / X_i \ge \tau), x \ge \tau$, from sample $S^{(n)}$ under nuisance pair(G_1, G_2) for specific number τ. In this article, we consider the problem of testing the composite hypothesis $H_0 :\ F \in \mathcal{F}$, where $\mathcal{F} = \{F(\cdot; \theta), \theta \in \Theta\}$- family of distribution depends on unknown parameter $\theta = (\theta_1, ..., \theta_s) \in \Theta$ and Θ- an open set in R^s. Consider two statistical tests for verify H_0 with a limit of chi-square distribution.

1 Generalized chi-square statistics

For build statistics of chi-square test we consider the nonparametric estimates of $1 - F_\tau(x)$ from [1]:

$$1 - F_{\tau n}(x) = \left[\frac{q_n(x)}{q_n(\tau)}\right]^{R_n(x;\tau)}, x \ge \tau,$$

where

$$R_n(x;\tau) = \Lambda_n(x;\tau)(L_n(x;\tau))^{-1}, \Lambda_n(x;\tau) = -\int_{[\tau;x]} (q_n(u))^{-1} dH_n^{(1)}(u),$$

$$L_n(x;\tau) = -\int_{[\tau;x]} (q_n(u))^{-1} dq_n(u), q_n(x) = G_{1n}(x) - H_n(x) + \frac{1}{n},$$

$$H_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(Z_i < x) = H_n^{(0)}(x) + H_n^{(1)}(x) + H_n^{(2)}(x),$$

$$H_n^{(m)}(x) = \frac{1}{n}\sum_{i=1}^{n} \delta_{mi} I(Z_i < x), \quad m = 0, 1, 2,$$

$$G_{1n}(x) = \exp\left\{-\int_{[x;\infty)} \left(H_n(u) + \frac{1}{n}\right)^{-1} dH_n^{(0)}(u)\right\},\ x \ge \tau.$$

In order to construct test statistics we introduce the conditions

(C1) Let d.f.-s F and G_1 continuous and the numbers τ, T such that, $\tau < T$ and

$$\inf_{\tau \le x \le T} P(Y_{1i} \le x \le X_i \wedge Y_{2i}) > 0;$$

(C2) Support $N_F = \{x : 0 < F(x;\theta) < 1\}$ independent on θ;

(C3) There is a density $f(x;\theta)$ with d.f.$F(x;\theta)$, it has continuous derivatives: $\frac{\partial^2 f(x;\theta)}{\partial\theta_i \partial\theta_j}$, $i, j = \overline{1, s}$ and $\int_{-\infty}^{+\infty} \left|\frac{\partial^2 f(x;\theta)}{\partial\theta_i \partial\theta_j}\right| dx < \infty$; $i, j = \overline{1, s}$;

(C4) Information matrix of Fisher $\mathrm{I}(\theta) = \|I_{ij}\|_{i,j=\overline{1,s}}$ is positive definite and continuous by θ, where

$$I_{ij}(\theta) = -\int_{-\infty}^{\infty} \frac{\partial^2 \log f(x;\theta)}{\partial\theta_i \partial\theta_j}(G_1(x) - G_2(x))dF(x;\theta) - \int_{-\infty}^{\infty} \frac{\partial^2 \log F(x;\theta)}{\partial\theta_i \partial\theta_j} F(x;\theta)dG_1(x) -$$

$$- \int_{-\infty}^{\infty} \frac{\partial^2 \log(1 - F(x;\theta))}{\partial\theta_i \partial\theta_j}(1 - F(x;\theta))dG_2(x);$$

(C5) There is a maximum likelihood estimate (MLE) $\widehat{\theta}_n = (\widehat{\theta}_{1n}, ..., \widehat{\theta}_{sn})$, for parameter$\theta = (\theta_1, ..., \theta_s)$, obtained by solving the system of equations

$$\frac{\partial \log p_n(\theta)}{\partial\theta_i} = 0,\ i = 1, ..., s,$$

where $p_n(\theta) = \prod_{i=1}^{n}(F(Z_i;\theta))^{\delta_{0i}}(f(Z_i;\theta))^{\delta_{1i}}(1-F(Z_i;\theta))^{\delta_{2i}}$ - the truncated likelihood function of the model. Moreover, the MLE $\widehat{\theta}_n$ can be represented by $n \to \infty$

$$n^{1/2}(\widehat{\theta}_n - \theta) = \mathrm{I}^{-1}(\theta)A_n(\theta) + o_p(1),$$

where $A_n(\theta) = n^{-1/2}\frac{\partial \log p_n(\theta)}{\partial\theta}$ is normalized contribution function.

We present the asymptotic properties of estimates $F_{\tau n}$ from [1]. We define a sequence of processes $\left\{V_n(x) = n^{1/2}(F_{\tau n}(x) - F_\tau(x)), x \geq \tau, n \geq 1\right\}$. For these processes the sequence of approximating processes is $\{M_n(x) = (1 - F_\tau(x))N_n(x)\}$, where

$$N_n(x) = \int_\tau^x \frac{(B_n(u) - \beta_n^*(u))dH^{(1)}(u)}{(G_1(u) - H(u-))^2} + \frac{B_n^{(1)}(x)}{G_1(x) - H(x-)} -$$

$$- \frac{B_n^{(1)}(\tau)}{G_1(\tau) - N(\tau-)} - \int_\tau^x \frac{B_n^{(1)}(u)d(G_1(u) - H(u-))}{(G_1(u) - H(u-))^2},$$

$$\beta_n^*(x) = -G_1(x)\left(\int_\tau^{+\infty} \frac{B_n(u)dH^{(0)}(u)}{H^2(u)} + \frac{B_n^{(0)}(x)}{H(x)} - \int_\tau^{+\infty} \frac{B_n^{(0)}(u)dH(u)}{H^2(u)}\right).$$

Here, for each n: $H(x) = P_{\theta_0}(Z_1 < x) = EH_n(x)$, $H^{(m)}(x) = P_{\theta_0}(Z_1 < x, \delta_{m1} = 1) = EH_n^{(m)}(x)$; $B_n(u) \stackrel{D}{=} B(H(u))$, $B_n^{(m)}(u) \stackrel{D}{=} B(H^{(m)}(u))$, $m = 0, 1, 2$ and $\{B(y), 0 \leq y \leq 1\}$ is process of a Brownian bridge. Note, that the processes $M_n(x)$ are linear functional of the Brownian bridge, and thus are Gaussian processes with zero mean. We present the following theorem from [1 , Theorem 2.1.2].

Theorem A. [1]. Under condition (C1) we have an approximation

$$P\left(\sup_{\tau \leq x \leq T} |V_n(x) - M_n(x)| > Rn^{-1/2}\log n\right) \leq Qn^{-\varepsilon},$$

where ε, $R = R(\varepsilon)$ and Q (absolute) positive constants.

Remark 1. In conditions of Theorem A for $\varepsilon > 1$ by lemma of Borel-Cantelli we have the strong approximation

$$\sup_{\tau \leq x \leq T} |V_n(x) - M_n(x)| \stackrel{\text{a.s.}}{=} O\left(n^{-1/2}\log n\right).$$

From here we have the weak convergence

$$V_n(x) \stackrel{D}{\Rightarrow} M(x) \text{ in } \mathrm{D}[\tau; T], \tag{1}$$

where $M_n(\cdot) \stackrel{D}{=} M(\cdot)$ for each n and Gaussian process $M(x)$ obtained from $M_n(x)$ by replacement of $B_n(u)$ and $B_n^{(m)}(u)$, $m = 0, 1, 2$ by the appropriate Brownian bridges with arguments $H(u)$ and $H^{(m)}(u), m = 0, 1, 2$ respectively.

We introduce the random processes $\varphi_n(x;\theta) = n^{1/2}(F_{\tau n}(x) - F_\tau(x;\theta))$, $\widehat{\varphi}_n(x) = \varphi_n(x;\widehat{\theta}_n)$. Let $\tau = x_o < x_1 < ... < x_{r-1} < x_r < T < \infty$ possible random partition for a given probability p_i, satisfying the equality $F(t_i;\widehat{\theta}_n) = p_i$. Consider a random vector $\widehat{\Phi}_n = (\widehat{\varphi}_n(x_1), ..., \widehat{\varphi}_n(x_r))^T$. The next result generalizes (1).

Theorem 1. Let for all $\theta \in \Theta$ the conditions (C1)-(C5) hold. Then the random process $\{\widehat{\varphi}_n(x), \tau \le x \le T\}$ converges weakly to the Gaussian process $\widehat{\varphi}_n(x)$ with zero mean and covariance with$\tau < x \le y < T$:

$$Cov_\theta\left(\widehat{\varphi}(x), \widehat{\varphi}(y)\right) = Cov_\theta\left(M(x), M(y)\right) - \left(\frac{\partial F(x;\theta)}{\partial \theta}\right)^T \mathrm{I}^{-1}(\theta) \frac{\partial F(y;\theta)}{\partial \theta},$$

where$\frac{\partial F(x;\theta)}{\partial \theta} = \left(\frac{\partial F(x;\theta)}{\partial \theta_1}, ..., \frac{\partial F(x;\theta)}{\partial \theta_s}\right)^s$.

Let $\widehat{M}$ and $\widehat{\Sigma}_0$ estimates of the matrices M and Σ_0, obtained by replacing θ on MLE $\widehat{\theta}_n$. The functions $G_1, H^{(m)}$ and H replaced by their nonparametric estimatesG_{1n}, $H_n^{(m)}$ and$H_n, m = 0, 1, 2$. Following the general principles of construction of chi-square statistics (see [5]), we consider the statistics

$$\Omega_n\left(\widehat{\theta}_n\right) = \widehat{\Phi}_n^T \widehat{\Sigma}_0 \widehat{\Phi}_n,$$

where $\widehat{\Phi}_n = (\widehat{\varphi}_n(x_1), ..., \widehat{\varphi}_n(x_r))^T$. Then we have

Theorem 2. Let the conditions (C1)-(C5) hold and $rang\left(\widehat{\Sigma}_0\right) = r$. Then

$$L\left(\Omega_n\left(\widehat{\theta}_n\right)/H_0\right) \underset{n\to\infty}{\to} K_r,$$

where K_r is chi-square distribution with degrees of freedom r.

2 Chi-square test based on the likelihood ratio statistics

First, θ is scalar parameter. Consider a simple hypothesis $H_0 : \theta = \theta_0$ against the composite alternative$H_1 : \theta \in \Theta_1$, where$\Theta = \{\theta_0\} \cup \Theta_1$. Let $\widehat{\theta}_n$ is MLE, satisfying the condition (C5) and consider LRS

$$L_n = \frac{p_n\left(\widehat{\theta}_n\right)}{p_n\left(\theta_0\right)}.$$

We also consider the condition

(C6) There is a third derivative on θ of density $f(x;\theta)$, exists the independent of θ function h_n such that

$$\left|\frac{d^3 \log p_n(\theta)}{d\theta^3}\right| \leq h_n; M_\theta h_n < \infty.$$

From the general theory of MLE (see [7]) follows that under conditions (C2) - (C4), (C6), there exists a unique consistent MLE $\widehat{\theta}_n$ and at $n \to \infty$

$$L\left(n^{1/2}\left(\widehat{\theta}_n - \theta_0\right)/H_0\right) \to N\left(0, \mathrm{I}^{-1}(\theta_0)\right).$$

Then by Taylor's formula

$$\log L_n = \log p_n(\theta_n) - \log p_n(\theta_0) = \frac{\partial \log p_n(\theta_0)}{\partial\theta}\left(\widehat{\theta}_n - \theta_0\right) +$$

$$+\frac{1}{2}\frac{\partial^2 \log p_n(\theta_0)}{\partial\theta^2}\left(\widehat{\theta}_n - \theta_0\right)^2$$

$$+\frac{1}{3!}\frac{\partial^3 \log p_n(\theta_*)}{\partial\theta^3}\left(\widehat{\theta}_n - \theta_0\right)^3 \tag{2}$$

and

$$\frac{\partial \log p_n(\widehat{\theta}_n)}{\partial\theta} - \frac{\partial \log p_n(\theta_0)}{\partial\theta} = \frac{\partial^2 \log p_n(\theta_{**})}{\partial\theta^2}\left(\widehat{\theta}_n - \theta_0\right) + \frac{1}{2}\frac{\partial^3 \log p_n(\theta_0)}{\partial\theta^3}\left(\widehat{\theta}_n - \theta_0\right)^2, \tag{3}$$

where $|\theta_* - \theta_0| \vee |\theta_{**} - \theta_0| \leq \left|\widehat{\theta}_n - \theta_0\right|$. Since, $\frac{\partial \log p_n(\theta_0)}{\partial\theta} = 0$, then substituting the expression for $\frac{\partial \log p_n(\widehat{\theta}_n)}{\partial\theta}$ from (3) into (2) we have

$$\log L_n = -\frac{1}{2}\frac{\partial^2 \log p_n(\theta_0)}{\partial\theta^2}\left(\widehat{\theta}_n - \theta_0\right)^2 + q_n,$$

where $q_n = o_p(1)$ at $n \to \infty$. Now, using the low of large numbers and central limit theorem, we find that under hypothesis H_0 at $n \to \infty$ statistics

$$-\frac{1}{n}\mathrm{I}^{-1}(\theta_0)\frac{\partial^2 \log p_n(\theta_0)}{\partial\theta^2}(n\mathrm{I}(\theta_0))\left(\widehat{\theta}_n - \theta_0\right)^2$$

have a chi-square distribution K_1 with one degree of freedom, i.e.

$$\mathcal{L}\left(2\log L_n/H_0\right) \to K_1. \tag{4}$$

Thus we have proved

Theorem 3. Under conditions (C2) - (C6) and the hypothesis H_0

$$2\log L_n \overset{D}{\Rightarrow} \chi^2_1. \tag{5}$$

Theorems 2 and 3 can be used to construct an asymptotic chi-square tests for the hypotheses H_0.

References

[1] Abdushukurov A. A. (2011). *Estimators of unknown distributions from incomplete observations and its properties*. LAMBERT Academic Publishing. 299 p. (In Russian).

[2] Abdushukurov A.A., Nurmukhamedova N.S. (2016) Approximation of likelihood ratio statistics in competing risks model under informative random censorship from both sides. Pakistan J. Statistics and Operation Research. V. XII. N. 1. p. 155-164.

[3] Habib M.G., Thomas D.R. (1986). Ch-square goodness-of-fit tests for randomly censored data. Ann. Statist. V.14. N2. p.759-765.

[4] Kaplan E.L., Meier P. (1958). Nonparametric estimation from incomplete observations. J.A.S.A. V.53. p.457-481.

[5] Voinov V., Nikulin M., Balakrishnan N. (2013). *Chi-Squared Goodness of Fit Tests with Application*. ELSEVIER. USA. 226 p.

[6] Voinov V., Naumov A., Pya N. (2003). Some recent advances in chi-squared testing. Proceeding of Internat. Conf. "Advances in Statistical Inferential Methods" : Almaty. Kazakhstan. p.233-247.

[7] Zacks Sh. (1971). *The theory oy statistical inference*. John Wiley & Sons, Inc. New York-London-Sydney-Toronto. 776p.

Empirical processes of independence in presence of estimated parameter

Leyla R. Kakadjanova
National University of Uzbekistan, Tashkent, Uzbekistan
e-mail: leyla_tvms@rambler.ru

Abstract

We consider the convergence of empirical processesof independence of random elements and events indexed by functions that depend on an estimated parameter and show that it can be replased by its natural limit.

Keywords: empirical process, independence, metric entropy, Glivenko-Cantelly and Donsker theorems.

Introduction

Folowing [1-2] consider a sequence of experiments in which observed data consist of pairs $\{(X_k, A_k), k \geq 1\}$, where X_k be independent and identically distributed random elements in a measurable space $(\mathfrak{X}, \mathfrak{B})$ with probability law $\mathbb{P}$ and A_k are events with common probability $p = \mathbb{P}(A_k) \in (0,1)$. Let $\delta_k = I(A_k)$ be an indicator of the event A_k. At the $n-$ th stage of experiments observed data is $\mathbb{S}^{(n)} = \{(X_k, \delta_k), 1 \leq k \leq n\}$. Each pair (X_k, δ_k) induced a statistical model with sample space $\mathfrak{X} \otimes \{0,1\}$ with σ-algebra $\mathcal{G}$ of sets $B \times D$ and distribution $\mathbb{Q}^*(\cdot)$ on $(\mathfrak{X} \otimes \{0,1\}, \mathcal{G})$:

$$\mathbb{Q}^*(B \times D) = \mathbb{P}(X_k \in B, \delta_k \in D), B \in \mathfrak{B}, D \subset \{0,1\}.$$

Our interests are focused on submeasures $\mathbb{Q}_m(B) = \mathbb{Q}^*(B \times \{m\})$, $m = 0,1$ and $\mathbb{Q}(B) = \mathbb{Q}_0(B) + \mathbb{Q}_1(B) = \mathbb{Q}^*(B \times \{0,1\})$, $B \in \mathfrak{B}$. From practical point of view, it is important to know the occurrence of hypothesis $\mathcal{H}$ about independence of random elements X_k and event A_k for each $k \geq 1$. In order to verify this we use the signed measure $\Lambda(B) = \mathbb{Q}_1(B) - p\mathbb{Q}(B)$, $B \in \mathfrak{B}$, where $p = \mathbb{Q}_1(\mathfrak{X})$. Then validity of $\mathcal{H}$ is equivalent to the equality $\Lambda(B) = 0$ for any $B \in \mathfrak{B}$. For constructing the test statistics for testing the hypothesis $\mathcal{H}$ we introduce the empirical analogue of measure $\Lambda(B)$ for any $B \in \mathfrak{B}$ as $\Lambda_n(B) = \mathbb{Q}_{1n}(B) - p_n\mathbb{Q}_n(B)$, where

$$p_n = \mathbb{Q}_{1n}(\mathfrak{X}), \mathbb{Q}_n(B) = \mathbb{Q}_{0n}(B) + \mathbb{Q}_{1n}(B),$$

$$\mathbb{Q}_{0n}(B) = \frac{1}{n}\sum_{k=1}^{n}(1-\delta_k)I(X_k \in B), \mathbb{Q}_{1n}(B) = \frac{1}{n}\sum_{k=1}^{n}\delta_k I(X_k \in B) \qquad (1)$$

- empirical counterparts of p, $\mathbb{Q}_(B)$ and $\mathbb{Q}_m(B), m = 0,1$, respectively. According to the strong law of large numbers (SLLN) for each $B \in \mathfrak{B}$ at $n \to \infty$, $\Lambda_n(B) \overset{\text{a.s.}}{\to} \Lambda(B)$ and consequently under validity of $\mathcal{H}$, $\Lambda_n(B) \overset{\text{a.s.}}{\to} 0$. Then naturally becomes investigation of limit behaviors of normalized signed empirical process $\{\chi_n = a_n(\Lambda_n(B) - \Lambda(B)), B \in \mathcal{J}\}$, indexed by certain class $\mathcal{J}$ of sets from $\mathfrak{B}$, where $\{a_n, n \geq 1\}$ is

a (possible random) sequence of positive numbers. In the previous works [1,2] we investigated the specially normalized empirical process of independence indexed by the class $\mathcal{F}$ of measurable functions f on $\mathfrak{X}$, which coincides with χ_n when $f = I(\cdot)$ is an indicator function.

1 Empirical Process of Independence Indexed by Class of Measurable Functions

For a signed measure G and class $\mathcal{F}$ of Borel measurable functions $f : \mathfrak{X} \to \mathbb{R}$ introduce the integral

$$Gf = \int\limits_{\mathfrak{X}} f dG, f \in \mathcal{F}.$$

Let's introduce the following $\mathcal{F}$-indexed extensions of empiricals (1) for $f \in \mathcal{F}$:

$$\mathbb{Q}_{0n} f = \frac{1}{n} \sum_{k=1}^{n} (1 - \delta_k) f(X_k), \;\; \mathbb{Q}_{1n} f = \frac{1}{n} \sum_{k=1}^{n} \delta_k f(X_k),$$

$$\mathbb{Q}_n f = \mathbb{Q}_{0n} f + \mathbb{Q}_{1n} f = \frac{1}{n} \sum_{k=1}^{n} f(X_k), \tag{2}$$

and $\Lambda_n f = \mathbb{Q}_{1n} f - p_n \mathbb{Q}_n f$, where $p_n = \mathbb{Q}_{1n} 1 = \mathbb{Q}_{1n}(\mathfrak{X}) = \frac{1}{n} \sum\limits_{k=1}^{n} \delta_k$. Observe that formulas (1) are special cases of (2) when $\mathcal{F} = \{I(B), B \in \mathcal{J}\}$. We define $\mathcal{F}$-indexed empirical process $G_n : \mathcal{F} \to \mathbb{R}$ as

$$f \mapsto G_n f = \sqrt{n} (\mathbb{Q}_n - \mathbb{Q}) f = n^{-1/2} \sum_{k=1}^{n} (f(X_k) - \mathbb{Q} f), f \in \mathcal{F}. \tag{3}$$

Here $G_n f = G_{0n} f + G_{1n} f$ with subempirical processes

$$G_{jn} f = \sqrt{n} (\mathbb{Q}_{jn} - \mathbb{Q}_j) f, \;\; j = 0, 1, f \in \mathcal{F}. \tag{4}$$

For a given f by SLLN and central limit theorem (CLT)

$$(a) \;\; \mathbb{Q}_n f \xrightarrow[n \to \infty]{a.s.} \mathbb{Q} f \;\text{ as }\; \mathbb{Q}|f| < \infty; \tag{5}$$

$$(b) \;\; G_n f \Rightarrow Gf \stackrel{d}{=} N\left(0, \sigma_{\mathbb{Q}}^2(f)\right), \;\; n \to \infty \;\text{ as }\; \mathbb{Q} f^2 < \infty, \tag{6}$$

where $\sigma_{\mathbb{Q}}^2(f) = \mathbb{Q}(f - \mathbb{Q} f)^2$.

Note that the uniformly variants for special classes $\mathcal{F}$ of measurable functions of statements (5) and (6) have a very solid theory (see, for example, [3,10]). There are different extensions of classical theorems of Glivenko-Cantelli and Donsker for $\mathcal{F}$-indexed empirical processes (3) under certain conditions on the set $\mathcal{F}$ of measurable functions. These conditions ensure the convergence of $n^{-1/2} \|G_n f\|_{\mathcal{F}} =$

$\sup\left\{n^{-1/2}\left|G_nf\right|, f\in\mathcal{F}\right\}$ either in probability or almost surely to zero. These classes $\mathcal{F}$ are called the weak or strong Glivenko-Cantelli classes respectively. Donsker-type theorems provide a general conditions on $\mathcal{F}$ in order to get a weak convergence

$$G_nf \Rightarrow Gf \quad \text{in} \quad l^{\infty}(\mathcal{F}), \tag{7}$$

where $l^{\infty}(\mathcal{F})$ is a space of all bounded functions $f:\mathfrak{X}\to\mathbb{R}$ equipped with the supremum-norm $\|.\|_{\mathcal{F}}$ (see, [8], p.81). Class $\mathcal{F}$ for which holds the convergence (7) is called a Donsker class. The limiting field $\{Gf,\ f\in\mathcal{F}\}$ in (7) is called a $\mathbb{Q}$-Brownian bridge. It is a tight Borel measurable element of $l^{\infty}(\mathcal{F})$ and is a Gaussian field with zero mean and covariance function

$$\operatorname{cov}(Gf,Gg)=\mathbb{Q}fg-\mathbb{Q}f\mathbb{Q}g,\quad f,g\in\mathcal{F}. \tag{8}$$

Remind that $\mathbb{Q}$-Brownian bridge $\{Gf, f\in\mathcal{F}\}$ can be represented in distribution though $\mathbb{Q}$-Brownian sheet $\{W(f), f\in\mathcal{F}\}$ with zero mean and covariance

$$\operatorname{cov}(W(f),W(g))=\mathbb{Q}fg,\quad f,g\in\mathcal{F}, \tag{9}$$

by distributional equality

$$Gf\overset{d}{=}W(f)-W(1)\,\mathbb{Q}f,\ \ f\in\mathcal{F}. \tag{10}$$

By SLLN under conditions $\mathbb{Q}_j|f|<\infty,\ j=0,1$ for given f:

$$\Lambda_nf\underset{n\to\infty}{\overset{a.s.}{\to}}\Lambda f\overset{under\ \mathcal{H}}{=}0. \tag{11}$$

Moreover, for a given f variable $\sqrt{n}(\Lambda_n-\Lambda)f$ being a linear functional of subempirical processes (4) under conditions $\mathbb{Q}_jf^2<\infty,\ j=0,1$, have a limiting normal distribution $N\left(0,\sigma^2_{\mathbb{Q}}(f)\right)$. In [1] the authors have proved uniform SLLN and CLT for a specially normalized empirical $\mathcal{F}$-indexed process

$$\left\{\Delta_nf=\left(\frac{n}{p_n(1-p_n)}\right)^{1/2}(\Lambda_n-\Lambda)f,\ \ f\in\mathcal{F}\right\}, \tag{12}$$

and showed that the limiting distribution is $\mathbb{Q}$-Brownian bridge $\{Gf,\ f\in\mathcal{F}\}$ with covariance (8).

2 Asymptotical results

Let $\mathcal{L}_q(\mathbb{Q})$ - the space of functions $f:\mathfrak{X}\to\mathbb{R}$ with the norm

$$\|f\|_{\mathbb{Q},q}=\left(\mathbb{Q}|f|^q\right)^{1/q}=\left\{\int\limits_{\mathfrak{X}}|f|^qd\mathbb{Q}\right\}^{1/q}.$$

To prove the $\mathcal{F}$ - uniform variants of theorems of Glivenko-Cantelli and Donsker we define the complexity or entropy of class $\mathcal{F}$. To determine the entropy it is necessary to define a concept of $\varepsilon-$ brackets. $\varepsilon-$ bracket in $\mathcal{L}_q(\mathbb{Q})$ is a pairs of functions $\varphi, \psi \in \mathcal{L}_q(\mathbb{Q})$ such that $\mathbb{Q}(\varphi(X) \leq \psi(X)) = 1$ and $\|\psi - \varphi\|_{\mathbb{Q},q} \leq \varepsilon$, i.e. $\mathbb{Q}(\psi - \varphi)^q \leq \varepsilon^q$. Function $f \in \mathcal{F}$ is in the (or covered by) bracket $[\varphi, \psi]$, if $\mathbb{Q}(\varphi(X) \leq f(X) \leq \psi(X)) = 1$. Note that the functions φ and ψ may not belong to the class $\mathcal{F}$, but they must have finite norms. Bracketing (or covering) number $N_{[]}(\varepsilon, \mathcal{F}, \mathcal{L}_q(\mathbb{Q}))$ is the minimum number of $\varepsilon-$ brackets in $\mathcal{L}_q(\mathbb{Q})$ needed to cover $\mathcal{F}$ (see [8,9]):

$$N_{[]}(\varepsilon, \mathcal{F}, \mathcal{L}_q(\mathbb{Q})) = \min \begin{cases} k: \text{ for some } f_1, ..., f_k \in \mathcal{L}_q(\mathbb{Q}), \\ \mathcal{F} \subset \bigcup\limits_{i,j} [f_i, f_j] : \|f_j - f_i\|_{\mathbb{Q},q} \leq \varepsilon. \end{cases}$$

Number $H_q(\varepsilon) = \log N_{[]}(\varepsilon, \mathcal{F}, \mathcal{L}_q(\mathbb{Q}))$ is called the metric entropy with bracketing of the class $\mathcal{F}$ in $\mathcal{L}_q(\mathbb{Q})$. Number $H_{jq}(\varepsilon) = \log N_{[]}(\varepsilon, \mathcal{F}, \mathcal{L}_q(\mathbb{Q}_j)), j = 0, 1$ denotes the metric entropy of a class $\mathcal{F}$ in $\mathcal{L}_q(\mathbb{Q}_j)$, $j = 0, 1$, respectively. To prove the weak convergence of $\mathcal{F}-$ indexed empirical processes (12) we introduce the integral of the metric entropy with bracketing as

$$J_{j[]}^{(q)}(\delta) = J_{j[]}(\delta; \mathcal{F}; \mathcal{L}_q(\mathbb{Q}_j)) = \int_0^\delta (H_{jq}(\varepsilon))^{1/2} d\varepsilon, j = 0, 1, \text{ for } 0 < \delta < 1.$$

Recall that numbers $N_{[]}(\cdot)$ converge to $+\infty$ at $\varepsilon \downarrow 0$. However, it necessary for Donsker's theorem that they converge not very fast to $+\infty$. This speed is measured by the integrals $J_{j[]}^{(q)}(\delta)$ (see [8,9]).

The following theorem shows validity of Glivenko-Cantelli type theorem for the process $\{\Delta_n f,\ f \in \mathcal{F}\}$. Here $*$ - means a.s. convergence by outer probability.

Theorem 1. [1] *Let the class $\mathcal{F}$ such that*

$$N_{[]}(\varepsilon, \mathcal{F}, \mathcal{L}_1(\mathbb{Q}_j)) < \infty, j = 0, 1. \tag{13}$$

Then under validity of the hypothesis $\mathcal{H}$ and at $n \to \infty$

$$\left\|n^{-1/2} \Delta_n f\right\|_{\mathcal{F}}^* \overset{a.s.}{\to} 0. \tag{14}$$

To prove the weak convergence of the process (12) to a Gaussian process, we first investigate the limiting properties of two-dimensional empirical field
$\{(\mathbb{A}_n f, \mathbb{A}_{1n} g),\ f, g \in \mathcal{F}\}$, where $\mathbb{A}_n f = = n^{1/2}(\mathbb{Q}_n - \mathbb{Q}) f$ and
$\mathbb{A}_{1n} g = n^{1/2}(\mathbb{Q}_{1n} - \mathbb{Q}_1) g$.

Theorem 2. [1] *Let the class $\mathcal{F}$ such that*

$$\mathcal{F} \subset \mathcal{L}_2(\mathbb{Q}_j) \text{ and } J_{j[]}^{(2)}(1) < \infty,\ j = 0, 1. \tag{15}$$

Then for $n \to \infty$ sequence $\{(\mathbb{A}_n f, \mathbb{A}_{1n} g),\ f, g \in \mathcal{F}\}$ of $\mathcal{F} \to \mathbb{R}^2$ maps weak converge in $l^\infty(\mathcal{F}) \times l^\infty(\mathcal{F})$ to the two-dimensional Gaussian field $\{(\mathbb{A} f, \mathbb{A}_1 g),\ f, g \in \mathcal{F}\}$ with zero mean and covariance structure for $f, g \in \mathcal{F}$:

$$\mathbb{E}(\mathbb{A} f \cdot \mathbb{A} g) = \mathbb{Q} f g - \mathbb{Q} f \mathbb{Q} g,$$

$$\mathbb{E}\,(\mathbb{A}_1 f \cdot \mathbb{A}_1 g) = \mathbb{Q}_1 f g - \mathbb{Q}_1 f \mathbb{Q}_1 g, \tag{16}$$
$$\mathbb{E}\,(\mathbb{A} f \cdot \mathbb{A}_1 g) = \mathbb{Q}_1 f g - \mathbb{Q} f \mathbb{Q}_1 g.$$

Remark. In the last formula of covariance in (16), in particular at $g \equiv 1$ we have $\mathbb{Q}_1 1 = p$ and

$$\mathbb{E}\,(\mathbb{A} f \cdot \mathbb{A}_1 1) = \mathbb{Q}_1 f - p\,\mathbb{Q} f,\ \ f \in \mathcal{F}. \tag{17}$$

Hence, when hypothesis $\mathcal{H}$ is valid, then the covariance (17) is equal to zero for all $f \in \mathcal{F}$. Thus, under the hypothesis $\mathcal{H}$, the Brownian bridge $\{\mathbb{A}\, f,\ \ f \in \mathcal{F}\}$ and r.v. $\mu_0 = \mathbb{A}_1 1$ with a normal distribution $\mathcal{N}\,(0, p\,(1-p))$ are independent.

In study of process (12) basic property of weak convergence to a $\mathbb{Q}-$ Brownian bridge is contained in the following statement.

Theorem 3. [1] *Under the conditions of Theorem 3.2 for* $n \to \infty$

$$\Delta_n f \Rightarrow \Delta f \text{ in } l^\infty(\mathcal{F}), \tag{18}$$

where $\{\Delta f,\ \ f \in \mathcal{F}\}$ *is a Gaussian field with zero mean and under validity of the hypothesis* $\mathcal{H}$*, it coincides in distribution with* $\mathbb{Q}-$ *Brownian bridge.*

Now we consider the convergence of empirical process of independence (12) when the class $\mathcal{F}$ of measurable functions $f_{\theta,\eta} : \mathfrak{X} \to \mathbb{R}$ indexed by sets Θ and $\mathcal{K}$: $\mathcal{F} = \{f_{\theta,\eta} : \theta \in \Theta, \eta \in \mathcal{K}\}$, where η is an estimated parameter. Let η_n be an estimator of η is a random element with values in $\mathcal{K}$ defined on the same probability space as $X_1, ..., X_n$ and $\eta_0 \in \mathcal{K}$ is a fixed element, which is limit in probability of the sequence η_n. In several applications it is interesting to prove that, as $n \to \infty$,

$$\sup_{\theta\in\Theta} |\Delta_n(f_{\theta,\eta_n} - f_{\theta,\eta_0})| \xrightarrow{p} 0. \tag{19}$$

We say that η_n is consistent for η_0, if

$$\sup_{\theta\in\Theta} \left|\Delta_n(f_{\theta,\eta_n} - f_{\theta,\eta_0})^2\right| \xrightarrow[n\to\infty]{p} 0. \tag{20}$$

Theorem 4. *If conditions (15) and (20) holds, then (19) is valid.*

The result (19) can be applicated to the estimaton of the functional $\theta \mapsto \Delta f_{\theta,\eta}$. When the parameter η is unknown, we can replace it by an estimator η_n an use the estimator $\Delta_n f_{\theta,\eta_n}$. This result helps to derive the limiting behaviour of this estimator by using the decomposition

$$(\Delta_n f_{\theta,\eta_n} - \Delta f_{\theta,\eta_0}) = \Delta_n(f_{\theta,\eta_n} - f_{\theta,\eta_0}) + \Delta_n f_{\theta,\eta_0} + \sqrt{\frac{n}{p_n(1-p_n)}}\Delta_n(f_{\theta,\eta_n} - f_{\theta,\eta_0}). \tag{21}$$

In right side of decomposition (21) the first term converges to zero by (19), the second term will converge to a Gaussian process by theorem 3 and the third term converges to zero by (20).

Acknowledgements

I would like to thank my supervisor, professor A. Abdushukurov, whose expertise was invaluable in the formulating of the research topic and methodology in particular.

References

[1] Abdushukurov A.A., Kakadjanova L.R. (2015). A class of special empirical process of independence. *J. of Siberian Federal Univ. Math. Phys.* Vol. **8**, N.2., pp. 1-9.

[2] Abdushukurov A.A., Kakadjanova L.R. (2018). Sequential empirical process of independence. *J. of Siberian Federal Univ. Math. Phys.* Vol. **11**, N.5., pp. 634-643.

[3] Alexander K.S. (1984). Probability inequalities for empirical processes and a law of the iterated logarithm. *Ann. Probab.* Vol. **12**, N.4., pp. 1041-1067.

[4] Dudley R.M. (1978). Central limit theorems for empirical measures. *Ann. Probab.* Vol. **6**, pp. 899-929.

[5] Gaensler P., Stute W. (1979). Empirical processes: a survey of results for independent and identically distributed random variables. *Ann. Probab.* Vol. **7**, N.2., pp. 193-243.

[6] Gine E., Zinn J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* Vol. **12**, N.4., pp. 929-989.

[7] Shorack G.R., Wellner J.A. (1986). *Empirical processes with applications to statisticsk.* John Wiley&Sons.

[8] Van der Vaart A.W., Wellner J.A. (1986). *Weak convergence and empirical processes.* Springer.

[9] Van der Vaart A.W. (1998). *Asymptotic Statistics.* Cambridge University Press, Cambridge.

[10] Van der Vaart A.W., Wellner J.A. (2007). IMS Lecture Notes-Monograph Series. Asymptotics: Particless, Processes and Inverse Problems. Vol. **55**, pp. 234-252.

Empirical likelihood confidence intervals for truncated integrals

D.G.Zakhidov, D.Kh.Iskandarov
Andijan branch of Tashkent agrar university, Andijan, Uzbekistan
e-mail: birinchi_dilshod@mail.ru, davlatbek1987@mail.ru

Abstract

In right random censoring model using limit behaviors of empirical likelihood function we propose Will-type asymptotic confidence intervals for truncated integrals.

Keywords: random censoring, relative-risk power estimator, truncated integral, confidence interval.

Let $X_1, X_2, ...$ (survival times) and $Y_1, Y_2, ...$ (censoring times) be two independent sequences of random variables (r.v.-s) on the real line with marginal distribution function (d.f.-s) F and G respectively. Under the right random censoring model, instead of observing X_i, we observe the pairs $(Z_i, \delta_i), i = 1, 2, ..., n$ where $Z_i = min(X_i, Y_i)$ and $\delta_i = I(X_i \leq Y_i)$ with $I(\cdot)$ the indicator function. Let H denote d.f. of Z_i. Then $H(t) = 1 - (1 - F(t))(1 - G(t))$. Let F and G are continuous. We are interested in constructing a nonparametric confidence interval for a integral functional of the form

$$\theta = \theta(F) = \int \varphi(t) dF(t)$$

where φ is some given Borel measurable function. Let F_n denote the Relative Risk Power estimator of F proposed [1] as

$$F_n(t) = 1 - [1 - H_n(t)]^{R_n(t)},\ t \in R, \tag{1}$$

where $H_n(t) = \frac{1}{n} \sum\limits_{i=1}^{n} I(Z_i \leq t)$ be empirical estimator of $H(t)$ and

$$R_n(t) = \sum_{i=1}^{n} \delta_i (Z_i \leq t) \left[1 - H_n(Z_i) + \frac{1}{n}\right]^{-1} \left\{ \sum_{i=1}^{n} I(Z_i \leq t) \left[1 - H_n(Z_i) + \frac{1}{n}\right]^{-1} \right\}^{-1}$$

is the relative-risk function estimator. Note that estimator (1) is a correct estimator of d.f. $F(t)$ than the Product-Limit estimator of Kaplan-Meier and Exponential-Hazard estimator of Altschuler-Breslow (see[2]). Since estimator (1) have same good properties such that representation as sum of independent and identically distributed (i.i.d) r.v.-s up to point $T < T_H = \inf\{t : H(t) = 1\}$, then instead of $\theta(F)$ we consider $\theta_T(F) = \int \varphi^*(t) dF(t)$ where $\varphi^*(t) = \varphi(t) I(t \leq T)$. We prove for plug-in estimator of $\theta_T(F)$ the asymptotic representation

$$\theta_T(F_n) = \int \varphi^*(t) dF_n(t) = \frac{1}{n} \sum_{i=1}^{n} U_i + O_p\left(n^{-\frac{1}{2}}\right) \tag{2}$$

where $U_i = U_i(F, G)$ are i.i.d. r.v.-s. with $EU_i = \theta_T(F)$. Following Owen's [3] idea we propose empirical likelihood confidence interval for truncated integral functional $\theta_T(F)$.

Let $V_i = \varphi * (Z_i) \triangle F_n(Z_i)$, where $\triangle F_n(Z_i) = F_n(Z_i) - F_n(Z_i-)$ is the jump of estimator (1) at point Z_i. We introduce estimated likelihood ratio function of θ_T by

$$R_n(\theta_T) = \max\{\prod_{i=1}^{n} np_i\}$$

subject to restrictions

$$n\sum_{i=1}^{n} p_i(V_i - \theta_T) = 0, \quad \sum_{i=1}^{n} p_i = 1.$$

Following by the method of Lagrange-multipliers, write

$$\mathcal{L}_m = \sum_{i=1}^{n} log(hp_i) - n\lambda\sum_{i=1}^{n} p_i(V_i - \theta_T) + \mu(\sum_{i=1}^{n} p_i - 1)$$

where λ and μ are Lagrange multipliers. Now setting to zero the partial derivative with respect to p_i, we have

$$\frac{\partial \mathcal{L}_m}{\partial p_i} = \frac{1}{p_i} - n\lambda(V_i - \theta_T) + \mu, \quad i = 1, ..., n.$$

From here, we obtain $\mu = -n$ and

$$p_i = [n(1 + \lambda_n(V_i - \theta_T))]^{-1}, \quad i = 1, ..., n,$$

where λ_n is the solution of the equation

$$\sum_{i=1}^{n} \frac{V_i - \theta_T}{1 + \lambda(V_i - \theta_T)} = 0. \tag{3}$$

Let

$$\sigma_U^2 = \lim_{n\to\infty} Var(n^{-1/2}\sum_{i=1}^{n} U_i) < \infty$$

and we define the log-likelihood function

$$L_n(\theta) = -2logR_n(\theta) = 2\sum_{i=1}^{n} log(1 + \lambda_n(V_i - \theta_T)).$$

Then under validity of representation (2), as $n \to \infty$

$$\sigma_U^{-2} Var(V_1) L_n(\theta_T) \xrightarrow{d} \chi_1^2, \tag{4}$$

where χ_1^2 is the chi-square r.v. with the degree of freedom one. Now using convergence (4) we can constructing the Wild-type asymptotical confidence intervals for truncated integral θ_T (for detals, see [3]).

References

[1] Abdushukurov A.A. (1998). Nonparametric estimation of distribution function based on relative risk function. *Commun. Statist.: Theory and Methods.* Vol. **27**, No.**8**, pp. 1991-2012.

[2] Abdushukurov A.A. (2009). Nonparametric estimation based on incomplete observations. *In: international Enciclopedia of Statistical Sciences.*(Prof. Miodrag Lovric ED.) Springer. 2009. Pt.**14**, pp. 962-964.

[3] Owen A. *Empirical likelihood.* Chapman and Hall /CRC. 2001.

Construction of basic durability model of drilling with using fuzzy regression models

A. A. Popov and V. S. Karmanov
Novosibirsk Technical State University, Novosibirsk, Russia
e-mail: a.popov@corp.nstu.ru, karmanov@corp.nstu.ru

Abstract

There are considered the questions of the application of mathematical models of metal cutting processes in the problem of optimizing machining modes. The class of models of resistance is determined and the algorithm for finding the optimal modes is formulated. For specific models, a practical solution is given.

Keywords: fuzzy regression model, cutting tool life, durability model of drilling.

Introduction

Determining the optimal modes is one of the main problem in the theory of metal cutting. To find the modes, it is necessary to use models that describe tool life. The method of finding the optimal modes depends on the type of model. The article describes an algorithm for constructing a basic model of cutting tool durability using a non-clear regression analysis apparatus using the example of a drilling operation.

1 Problem Statement

The strength of the drill is characterized by the total length of the holes [1, 2, 3, 4], drilled by the tool before blunting - L, mm; or the operating time of the tool before regrinding - T, min. Tool durability depends on two factors - feed rate per revolution S, mm/rev (or minute feed speed S_m, mm/min) and rotational speed n, rpm (or rotational speed V, m/min).

These values are related by the following ratios:

$$L = S_m T, \tag{1}$$

$$S_m = Sn, \tag{2}$$

$$V = \frac{\pi d n}{1000}, \tag{3}$$

where d is the diameter of the hole, mm.

Thus, a general view of the resistance model:

$$L = f(S, n) \text{ or } T = f(S, n). \tag{4}$$

The parameters of the model (4) are estimated from the data of the persistence experiment. Next, the resulting model is used to find the optimal cutting conditions. For example, the criterion of optimality can be selected criterion of minimum costs [2, 3, 4, 5, 6, 7]

$$Q(n,S) = \frac{C}{L(n,S)} + \frac{D}{Sn} + E, \tag{5}$$

where C, D, E – some economic parameters. These criteria include the criterion of minimum costs, minimum cost (single unit of production, production line, factory, industry, etc., which is determined by the composition of the parameters C, D, E).

2 Fuzzy regression models

$$y = f^T(x)\theta + e = \sum_{l}^{m} f_l(x)\theta_l + e, \tag{6}$$

where y is the resulting attribute (response variable, random dependent variable); $f^T(x) = (f_1(x), f_2(x), ..., f_m(x))$ a given vector function of an independent variable $x = (x_1, ..., x_k)^T$, which may vary in the region $\tilde{X}$; $\theta = (\theta_1, ..., \theta_m)^T$ - unknown parameters that need to be determined from the results of experiments (measurements); e is an observation error. Due to the complexity of the modeled object and the influence of unaccounted factors, it is not always possible to unambiguously determine the structure of the vector $f(x)$. It can often be observed that in different parts of the regressors definition domain different models can be more adequate. One of the most effective modeling methods in this case is the concept of fuzzy systems [8, 9, 10, 11, 12, 13].

Fuzzy regression models will be specified by means of a regression tree. Let $x_1, x_2, ..., x_k$ be linguistic variables. Their values are determined by fuzzy sets $A, B, ..., \Gamma$, and the degree of intensity of the manifestation of the value will be set as the value of the functions of belonging. The branches of the decision tree have the form [14].

$$\Pi_{ij..l} : \text{If } (x_1 \text{ is } A_i) \wedge (x_2 \text{ is } B_j) \wedge ... \wedge (x_k \text{ is } \Gamma_l) \text{ then } y'_{ij..l} = \eta + \alpha_i + \beta_j + ... + \gamma_l. \tag{7}$$

The truth of statements $(x_1 \text{ is } A_i), (x_2 \text{ is } B_j), ..., (x_k \text{ is } \Gamma_l)$ is determined by the values of the corresponding membership functions $\mu_{A_i} \in [0,1], \mu_{B_j} \in [0,1], ..., \mu_{\Gamma_l} \in [0,1]$. The degree of truth of a statement $\prod_{ij...l}$ will be denoted as $\mu(y'_{ij..l})$ and calculated as $\mu(y'_i j..l) = \mu_{A_i}\mu_{B_j}..\mu_{\Gamma_l}$. Regarding the assignment of values $\mu_{A_i}, \mu_{B_j}, ..., \mu_{\Gamma_l}$, we introduce the requirement that for each observation the following conditions are met:

$$\sum_{i=1}^{I} \mu_{A_i} = 1, \sum_{j=1}^{J} \mu_{B_j} = 1, ..., \sum_{l=1}^{L} \mu_{\Gamma_l} = 1;$$

$$\mu_{A_i} \in [0,1], i = \overline{1,I}, \mu_{B_j} \in [0,1], j = \overline{1,J}, ..., \mu_{\Gamma_l} \in [0,1], l = \overline{1,L}. \quad (8)$$

The procedure of dephasing is carried out according to the scheme:

$$y_{ij..l} = \frac{\sum \mu(y'_{ij..l}) y'_{ij..l}}{\sum \mu(y'_{ij..l})}. \quad (9)$$

Taking into account (8), the decision tree (7) can be represented as an observation model

$$y_{ij..l} = \eta + \sum_{i=1}^{I} \mu_{A_i}\alpha_i + \sum_{j=1}^{J} \mu_{B_j}\beta_j + ... + \sum_{l=1}^{L} \mu_{\Gamma_l}\gamma_l + e_{ij..l}. \quad (10)$$

After finding the estimate $\hat{\theta}^T = (\hat{\eta}, \hat{\alpha}_1, ..., \hat{\alpha}_I, \hat{\beta}_1, ..., \hat{\beta}_J, \hat{\gamma}_1, ..., \hat{\gamma}_L,)$ for the parameter vector of parameters θ, the decision tree can be fixed as

$$\hat{\Pi}_{ij..l} : \text{If } (x_1 \text{ is } A_i) \wedge (x_2 \text{ is } B_j) \wedge ... \wedge (x_k \text{ is } \Gamma_l) \text{ then } \hat{y}'_{ij..l} = \hat{\eta} + \hat{\alpha}_i + \hat{\beta}_j + ... + \hat{\gamma}_l. \quad (11)$$

It can also be represented as a convolution.

$$\hat{y} = \hat{\eta} + \sum_{i=1}^{I} \mu_{A_i}\hat{\alpha}_i + \sum_{j=1}^{J} \mu_{B_j}\hat{\beta}_j + \sum_{l=1}^{L} \mu_{\Gamma_l}\hat{\gamma}_l. \quad (12)$$

The considered techniques for representing decision trees in the form of statements (11) and (12) can be extended to the case of using explanatory variables, measured on a quantitative scale. To simplify the presentation, we consider a particular case when the number of input factors is two. We break up the scope of the quantitive variables x_1, x_2 into fuzzy partitions, which, as before, for the first factor we will denote as $A_1, A_2, ..., A_I$ with the corresponding membership functions $\mu_{1i} \in [0,1]$, $i = \overline{1,I}$. Similarly, for the factor x_2, these will be partitions $B_1, B_2, ..., B_J$ with accessory functions $\mu_{2j} \in [0,1]$, $j = \overline{1,J}$. We will proceed from the fact that at individual sufficiently wide intervals of the action of quantitative factors, the behavior of the system response can be described by a linear dependence. In this case, the complexity of the tree can be reduced by trying to replace the representation of the response in the leaves of the tree, for example, by its linear dependence on input factors. The decision tree from two factors in this case will consist of branches of the form

$$\Pi_{ij} : \text{ If } (x_1 \text{ is } A_i) \wedge (x_2 \text{ is } B_j) \text{ then}$$
$$y'_{ij} = \theta_0 + \theta_{01i} + \theta_{02j} + (\theta_1 + \theta_{11i} + \theta_{12j})x_1 + (\theta_2 + \theta_{21i} + \theta_{22j})x_2. \quad (13)$$

Here, a part of the terms, namely $\theta_0 + \theta_1 x_1 + \theta_2 x_2$, is included in each branch of the tree and determines the overall linear dependence of the response on input factors throughout the entire domain of their definition without taking it into account for splitting it into partitions. Taking into account (8), decision tree (13) can be represented as an observation model

$$y_{ijl} = \theta_0 + \sum_{i=1}^{I} \mu_{1i}\theta_{01i} + \sum_{j=1}^{J} \mu_{2j}\theta_{02j} + (\theta_1 + \sum_{i=1}^{I} \mu_{1i}\theta_{11i} + \sum_{j=1}^{J} \mu_{2j}\theta_{12j})x_1 +$$
$$+(\theta_2 + \sum_{i=1}^{I} \mu_{1i}\theta_{21i} + \sum_{j=1}^{J} \mu_{2j}\theta_{22j})x_2 + e_{jil}. \quad (14)$$

After estimating the parameters θ, the decision tree in the form of convolution takes the form

$$\hat{y} = \hat{\theta}_0 + \sum_{i=1}^{I} \mu_{1i}\hat{\theta}_{01i} + \sum_{j=1}^{J} \mu_{2j}\hat{\theta}_{02j} + (\hat{\theta}_1 + \sum_{i=1}^{I} \mu_{1i}\hat{\theta}_{11i} + \sum_{j=1}^{J} \mu_{2j}\hat{\theta}_{12j})x_1 +$$
$$+(\hat{\theta}_2 + \sum_{i=1}^{I} \mu_{1i}\hat{\theta}_{21i} + \sum_{j=1}^{J} \mu_{2j}\hat{\theta}_{22j})x_2. \quad (15)$$

To ensure the identifiability of model (14), we will carry out its reduction by removing a series of regressors from it. For example, you can remove regressors μ_{1I}, μ_{2J} from the model, as well as $\mu_{1I}x_1, \mu_{1I}x_2, \mu_{2J}x_1, \mu_{2J}x_2$. The rationale for this method of identifying the model can be found in [15, 16, 17, 18].

When using fuzzy regression models, it is necessary to make assumptions about the number, form and location of fuzzy partitions for each factor. In this paper, we restrict ourselves to the consideration of linear and quadratic local models. At the same time, the possibility of using local models of increased complexity is limited primarily by the implemented experiment plan. In the problem considered in the work, the implemented plan is a complete factorial experiment at 5 levels for two factors — a total of 25 different points, not counting repeated observations. This allows to use a quadratic polynomial of two factors as a local model when dividing the domain of definition of the factors into two partitions. As for the form of the function of belonging, for the considered problems we will use trapezoidal ones. To the category "form of partitions" also should be referred the coordinates of the intersection points of the neighboring fuzzy partitions and the width of their intersection. By virtue of the symmetry of the domain of definition of factors with respect to zero, fuzzy partitions can also be placed symmetrically with respect to a zero value. The coordinate of the point of intersection of the partitions will be denoted as $\overline{x_\mu}$. We denote the half width of the intersection as $\Delta\mu$. The width of the partition cross section directly affects the smoothness of the transition of the regression dependence from one local model to another. When solving a specific task of restoring dependency, the parameters $\overline{x_\mu}$ and $\Delta\mu$ can be customized.

3 Solution of the practical task

The proposed method of calculation was used to determine the optimal modes (n^*, S^*) of drilling IXI8H9T stainless steel with a drill of $\varnothing 4.2\ mm$ using coolant.

For the experiment, was developed and upgraded the bench drilling machine for smooth regulation in a wide range of n and S.

In tab. 1 presents the data of the durability experiment [3, 4]. The experiments were carried out according to the plan of a full factorial experiment of type 5^2.

Table 1: Data of the stability experiment L, mm (drill ∅4.2, drilling depth 2d, drill reach 10d)

S, mm/rev	n, rpm				
	750	1098	1447	1795	2145
0.0280	570	1430	3600	1400	430
	390	1370	1800	1200	250
0.0450	5560	8300	4700	4000	700
	8500	6300	5700	3000	1100
0.0621	4040	5800	6130	3330	590
	5640	7800	4230	4070	810
0.0790	3150	3420	2760	1350	470
	3850	4180	3800	1650	690
0.0962	1910	130	100	30	9
	3170	150	140	50	11

Previously in [4, 5, 7], the logarithmic quadratic model was used as the base model describing the data of the stochastic experiments (see tab. 1), where the response values L were subjected to logarithms. In this case, logarithm was used to reduce the spread of response. When using nonlinear transformation of the response, the requirement for accuracy of approximation of such a response by one or another regression model significantly increases. The scatter of the response scale (see tab. 1) also indicates that the experiment was conducted on fairly wide ranges of factors. On the wide ranges of input factors, it is often possible to observe the model drift, when in different parts of the factor definition area the nature of the dependence of the response changes.

An important point in building a workable model of the process being studied is to decide on the choice of the optimal complexity model, which would not have the effect of retraining. It is called the effect of retraining if a model is tuned to the description of the training data and has poor predictive properties. When choosing a model, one should be guided by the external quality criteria [16]. Going over models of varying complexity, we choose in this case a model for which the external criterion is minimal. We will rely in this paper on two such criteria. Suppose that the sample of observations is divided into two parts A and B. In the methods of structural optimization, the following external criteria for model selection are widely used [21-23]:

regularity criterion:

$$\Delta^2(B) = \Delta^2(B/A) = ||y_B - X_B\hat{\theta}_A||^2, \tag{16}$$

where X_B is the matrix of observations on the part B; θ_A – estimates of the model parameters obtained by the sample A.

External criteria also includes the criterion of "sliding control" (CV - cross validation):

$$\Delta^2_{ck} = \sum_i (y_i - f^T(x_i)\hat{\theta}_{(i)})^2, \tag{17}$$

where $\theta_{(i)}$ is the parameter estimate for the full sample with excluded i observation.

Consider several classes of models and criteria values that are achieved on them. The results are shown in tab. 2, where the residual sum of squares for the given models is denoted by ε^2. Structures of optimal models in their class of models were selected by a minimum $\Delta^2(B)$ of the regularity criterion, two other criteria Δ^2_{ck} and ε^2 were additionally fixed.

An analysis of table 2 shows that the use of ordinary polynomials of the second and third degree does not significantly improve the quality of the approximation of experimental data.

Table 2: Indicators of quality models of optimal complexity

Model	Number of parameters, s	$\Delta^2(B)$	Δ^2_{ck}	ε^2
Linear	3	47	2.07	89.4
Linear with interactions	3	47	2.07	89.4
Quadratic	4	16.3	0.78	31.3
Cubic	10	6.43	0.304	9.07
Fuzzy linear, $\overline{x_\mu} = 0$, $\Delta\mu = 0.5$	9	15.3	0.79	25.8
Fuzzy linear with interactions, $\overline{x_\mu} = 0$, $\Delta\mu = 0.5$	10	10.9	0.53	17.4
Fuzzy quadratic, $\overline{x_\mu} = 0$, $\Delta\mu = 0.5$	12	5.45	0.229	7.15
Fuzzy quadratic with 2 batches on S, $\overline{x_\mu} = 0$, $\Delta\mu = 0.5$	10	5.34	0.242	7.05

For the models presented in tab. 2, it is possible to test the hypothesis of their adequacy by calculating the statistics $F = \hat{\sigma}^2_{LF}/\hat{\sigma}^2_e$, where $\hat{\sigma}^2_{LF} = \varepsilon^2/(N - s)$ is the estimate of the variance of observations obtained from the model; $\hat{\sigma}^2_e$ – estimation of variance by repeated observations. In our case $\hat{\sigma}^2_e = 0.0612$. Characteristically that none of the models presented in tab. 2 according to the criterion F is recognized as adequate. In order for this to happen, it is necessary to ensure the accuracy of the approximation with the estimated variance of the model $\hat{\sigma}^2_{LF} \simeq 0.1$. This can be

done if the tuning functions of the membership in the parameters $\overline{x_\mu}$ and $\Delta\mu$. The results presented in tab. 2 suggest that the optimal complexity of the model may be in the range of 10–12 parameters. The result was a fuzzy quadratic model with $\overline{x_\mu} = -0.22$, $\Delta\mu = 0.73$. Its optimal structure has 12 parameters and provides the following quality indicators: $\Delta^2(B) = 3.66$, $\Delta^2_{ck} = 0.1427$, $\varepsilon^2 = 3.94$. The differences in the behavior of the restored dependencies by quadratic and fuzzy quadratic models can be traced in Figures 1 and 2, where the label "y" denotes the observed response values.

The feed rate S varies in the range [–1, +1] in normalized units, which corresponds to the range [0.0280; 0.0962].

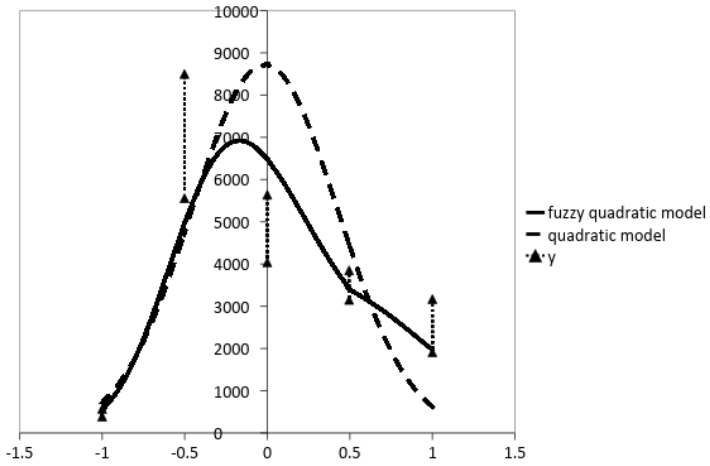


Figure 1: Sections for a fuzzy quadratic and quadratic model, n=750.

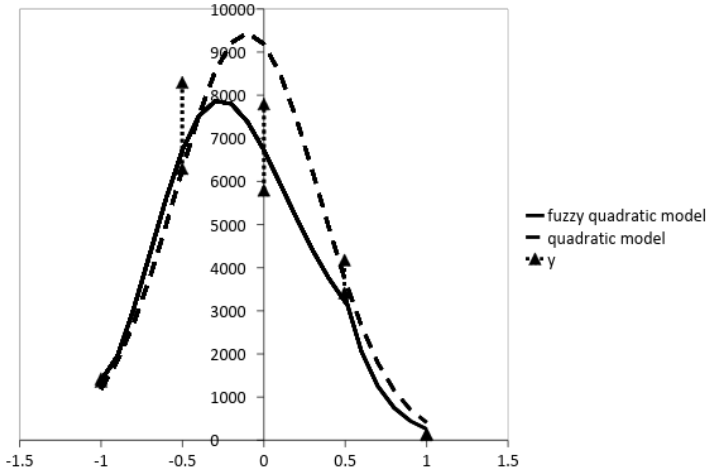


Figure 2: Sections for a fuzzy quadratic and quadratic model, n=1098.

Conclusion

A comparative analysis of several types of resistance drilling models has been carried out. Fuzzy quadratic model is proposed as an adequate model proposed for practical use.

References

[1] Smagin G.I., Karmanov V.S. (1999). Optimization of modes - a way to save costs. *Siberian Instrument*. Vol. **3**.

[2] Smagin G.I., Karmanov V.S. (2005). CHARACTERISTICS AND CRITERIA OF INSISTANCE RESISTANCE IN BLACK AND CLEANING BY TURNING *Metal processing (technology, equipment, tools)*. Vol. **4**, pp. 34-35.

[3] Karmanov V.S., Smagin G.I. (2006). CORRECTION OF THE EXTRAPOLATION AREA OF THE X-DRESSING STRUCTURE MODEL IN THE NORMALIZATION OF THE CUTTING MODES OF THE HARD-PROCESSED MATERIALS (ON THE EXAMPLE OF DRILLING) *Processing of metals (technology, equipment, tools)*. Vol. **2**.

[4] Karmanov V.S. (2006). RESEARCH OF MATHEMATICAL MODELS OF STABILITY OF A REHABLING TOOL *Scientific Bulletin of Novosibirsk State Technical University*. Vol. **2**, pp. 55-64.

[5] Smagin G.I., Karmanov V.S. (2009). APPLICATION OF THE METHOD OF CHARACTERISTIC LINES AND CHARACTERISTIC SURFACES IN NORMING THE CUTTING MODES *Processing of metals (technology, equipment, tools)*. Vol. **1**, pp. 16-19.

[6] Chimitova E., M., Karmenov V., Smagin G. (2015). THE CONSTRUCTION OF ACCELERATED LIFE MODEL FOR RELIABILITY OF A CUTTING TOOL *In collection: APPLIED METHODS OF STATISTICAL ANALYSIS. NONPARAMETRIC APPROACH proceedings of the international workshop*. pp. 220-226.

[7] G. I. Smagin, V. S. Karmanov, I. V. Fedin (2015). Using the basic model of the drilling process for rationing cutting conditions *Metal processing: technology, equipment, tools*. Vol. **4 (69)**, pp. 6-17.

[8] Takagi T., Sugeno M. (1985). Fuzzy Identification of Systems and Controls. IEEE Trans. on Systems, Man and Cybernetics. V. 15. No. 1. pp. 116–132.

[9] R. Babuska (1998). *Fuzzy Modeling for Control*. London Boston: Kluwer Academic Publishers.

[10] John H. Lilly (2010). *Fuzzy Control and Identification*. Wiley.

[11] A. Pegat (2013). *Fuzzy modeling and control. Tr. from English -2 ed.* Binom. Moscow.

[12] Popov A. A. (2000). Regression modeling based on fuzzy rules *Sb. scientific tr. NSTU.* Vol. **2 (19)**, pp. 49-57.

[13] Popov A.A., Bykhanov K.V. (2005). Modeling volatility of the time series using the GARCH models *Proceedings - 9th Russian-Korean International Symposium on Science and Technology, KO-RUS-2005 sponsors: Novosibirsk State Technical University.* pp. 687-692.

[14] Popov A. A. (2009). Building decision trees to predict a quantitative sign on a class of logical functions of linguistic variables *Scientific Bulletin of NSTU.* Vol. **3 (36)**, pp. 77–86.

[15] Popov A. A. (1996). Designing discrete and continuous-discrete regression-type models *Collected scientific works of NSTU.* Vol. **1**, pp. 21–30.

[16] Popov A.A. (2013). *Optimal planning of the experiment in the tasks of structural and parametric identification of models of multifactor systems: monograph.* Publishing house NSTU, Novosibirsk.

[17] Popov A.A., Lyakh K.N. (2003). Identifiability of models of soft analysis of variance *Collection of scientific works of NSTU.* Vol. **1 (31)**, pp. 79-84.

[18] Popov A.A., Lyakh K.N. (2003). Analysis of linear models of soft analysis of variance *Collection of scientific works of NSTU.* Vol. **1 (31)**, pp. 85-90.

The Wiener degradation model in the analysis of the laser module ILPN-134[1]

Evgeniya S. Chetvertakova, Ekaterina V. Chimitova,
Evgeniya A. Osintseva, Roman V. Snetkov
Novosibirsk State Technical University, Novosibirsk, Russia
e-mail: chimitova@corp.nstu.ru, evgenia.chetvertakova@gmail.com

Abstract

In this paper, we consider the application of the Wiener degradation model in the analysis of the laser module ILPN-134 degradation data. The Wiener degradation model is based on the assumption that degradation increments are independent normally distributed random variables. It has been shown that an appropriate model for the considered data is the Wiener degradation model with random effects, which takes into account the unit-to-unit variability. The constructed model can be used for estimation of the reliability indicators, such as the probability of the non-failure operation during some period of time.

Keywords: degradation process, Wiener degradation model, random effect, maximum likelihood estimation, reliability.

Introduction

The reliability analysis is an important stage for the quality assessment of technical systems and devices. The mathematical apparatus of reliability analysis includes methods for the construction of a statistical model describing the lifetime distribution. There are two approaches to construction of statistical reliability models: the first one uses only the information about observed failures, but the second approach considers all values of some indicator (index) characterizing the degradation process. In that way, statistical degradation models use more information about tested items to estimate the reliability than the models based on the samples of failure time data. So, if there is the possibility to observe the degradation paths until failure instead of only the time to failure, it will allow to obtain more accurate estimates of probability of the non-failure operation during some period of time.

The most popular statistical degradation models for the analysis of real data is the gamma and Wiener degradation models. For example, in [6], the gamma model is considered to analyze the wear of car tires depending on various stress factors, and in [3] and [8], the authors use the gamma degradation model to describe the aging of automobile brake pads. However, the gamma model cannot be applied in the cases, when values of degradation increments are not positive. In such cases, it is necessary to construct the Wiener degradation model.

If unit-to-unit variability is observed in investigated data, a "random-effect" degradation model should be considered. For such models, the distribution of random

[1]This work is supported by the Russian Ministry of Education and Science (project 1.1009.2017/4.6)

parameter is taken into account. For example, in [7], the "random-effect" Wiener degradation model with random scale parameter from the gamma distribution is offered. In this case, the number of unknown parameters of the "random-effect" model is larger than for the "fixed-effect" model, and as a result, the accuracy of parameter estimation for the "random-effect" model may decrease. On the other hand, if the unit-to-unit variability is rather large, then the "fixed-effect" model is not appropriate [1].

In this paper, we construct the "random-effect" Wiener degradation model for the reliability analysis of the semiconducting laser module ILNP-134 [9].

1 The problem formulation for the reliability analysis of the laser module ILPN-134

The reliability analysis of the semiconducting laser module ILPN-134 was described in [9]. 15 lasers were divided into three groups of 5 items and tested under the temperature of 70°C, 80°C and 90°C in each group, respectively. It was necessary to maintain a consumption current corresponding to a radiation power of 3 mW during the 8500 hours of accelerated tests. During the experiment, two items were excluded from the consideration because of the insufficient adjustment on maximum of the output optical power.

The value of the current (degradation index) was measured every 100 ± 20 h. An item fails when the current value is 20% higher than the initial value.

The degradation paths for different groups of the experiment are shown on Figures 1–3.

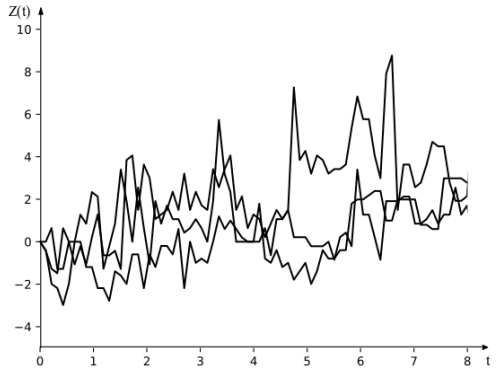


Figure 1: Observed degradation paths under 70°C

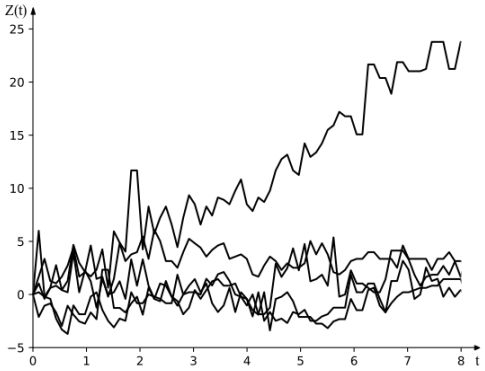


Figure 2: Observed degradation paths under 80°C

As can be seen from the figures, the typical degradation process is observed for the first two tested groups which relates to the aging of the laser diodes. However, the other nature of the degradation with the premature failures is detected for two lasers tested under the stress of 90°C. The detailed research of these devices showed

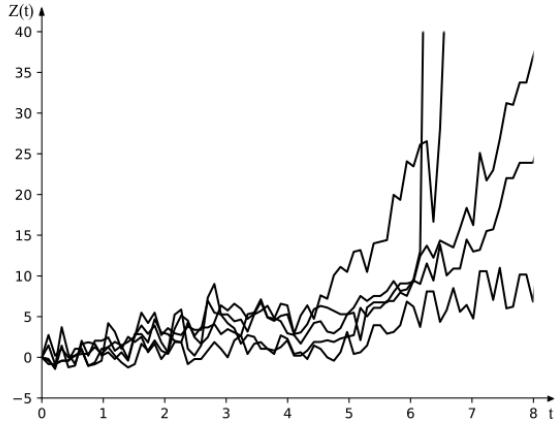


Figure 3: Observed degradation paths
under 90°C

that the reason for this behavior is connected to the optical fiber offset relative to the laser diode.

In [9], the authors carried out the reliability analysis basing only on the failure time values which were obtained by the approximation of degradation paths with the exponential function for each tested laser. In this paper, we construct a statistical model using all the values of the degradation index. Obtained increments of the degradation index are positive and negative, so it is reasonable to use the Wiener degradation model for the further analysis.

2 The Wiener degradation models

Let us assume that the observed stochastic process $Z(t)$ is a stochastic process with independent increments, and $Z(0) = 0$. For the Wiener degradation model, increments $\Delta Z(t) = Z(t + \Delta t) - Z(t)$ have the normal distribution with the probability density function

$$f_{Norm}(u; \theta_1, \theta_2) = \frac{1}{\sqrt{2\pi}\theta_2} e^{-\frac{(u-\theta_1)^2}{2\theta_2^2}}, \tag{1}$$

where $\theta_1 = \mu(\nu(t + \Delta t) - \nu(t))$ is the shift parameter and $\theta_2 = \sigma$ is the scale parameter, $\nu(t)$ is a positive increasing function.

Let the degradation process $Z(t)$ is observed under a constant in time stress (covariate) x , the range of values of which is defined by the conditions of the experiment. There are various ways to parameterize the dependence of the degradation path on covariates. Here, we assume that the covariate x influences the degradation paths as in the accelerated failure time model [5]:

$$Z_x(t) = Z\left(\frac{t}{r(x; \beta)}\right), \tag{2}$$

where $r(x;\beta)$ is the positive covariate function, β is the vector or scalar regression parameter.

Denote the mathematical expectation of degradation process $Z_x(t)$ by

$$M\left(Z_x(t)\right) = m_x(t) = \mu\nu_x(t) = \mu\nu\left(\frac{t}{r(x;\beta)};\gamma\right), \tag{3}$$

where $m_x(t)$ is the trend function of the degradation process, γ is the vector or scalar trend parameter.

The time to failure, which depends on covariate x , is defined as:

$$T_x = \sup\{t : Z_x(t) < z_0\}, \tag{4}$$

where z_0 is the critical value of the degradation index. Then, the reliability function for the Wiener degradation model can be written as:

$$S_x\left(t\right) = P\left\{T_x > t\right\} = P\left\{Z_x\left(t\right) < z_0\right\} = \Phi\left(\frac{z_0 - m_x\left(t\right)}{\sigma}\right), \tag{5}$$

where $\Phi(\cdot)$ is the standard normal distribution function.

To take into account unit-to-unit variability, the random effect can be included into the model by considering the parameter μ as a random variable from truncated normal distribution with the density function [1]

$$f_{trunc}(t;\delta,\alpha) = \frac{f_{Norm}(t;\delta,\alpha)}{1 - F_{Norm}(0;\delta,\alpha)}, \tag{6}$$

where δ is the shift parameter and α is the scale parameter.

Then, the marginal density function for $Z_x(t)$ in the case of Wiener degradation model with random effects is equal to [2]:

$$f_{Z_x(t)}\left(u;\nu_x(t),\sigma,\delta,\alpha\right) = \int_0^\infty f_{Norm}\left(u;\omega\nu_x(t),\sigma\right) f_{trunc}\left(\omega;\delta,\alpha\right) d\omega \tag{7}$$

In this case, the reliability function can be written as

$$S_x(t) = P\{T_x > t\} = P\{Z_x(t) < z_0\} = \int_0^{z_0} f_{Z_x(t)}\left(u;\nu_x(t),\sigma,\delta,\alpha\right) du \tag{8}$$

Let the realization of stochastic process $Z(t)$ for the i-th item under the value of covariate $x = x^i$ is denoted as

$$Z^i = \left\{(0, Z_0^i = 0),\ (t_1^i, Z_1^i), ..., (t_{k_i}^i, Z_{k_i}^i)\right\}, i = \overline{1,n}, \tag{9}$$

where k_i is the number of time moments, in which the degradation index was measured. Then, the sample of independent degradation index increments with covariates can be written as:

$$\mathbf{X}_n = \left\{\left(X_j^i = Z_j^i - Z_{j-1}^i, x^i\right),\ i = \overline{1,n},\ j = \overline{1,k_i}\right\} \tag{10}$$

Maximum likelihood estimates of parameters σ, γ and β of the "fixed-effect" Wiener degradation model are calculated by maximization of the likelihood function:

$$L(\mathbf{X}_n) = \prod_{i=1}^{n} \prod_{j=1}^{k_i} f_{Norm}\left(X_j^i; \mu\nu_{x^i}(t;\gamma,\beta), \sigma\right). \tag{11}$$

If $Z^i(t), i = \overline{1,n}$ are the Wiener degradation processes with random effects, then the likelihood function can be written as the multiplication of the joint density functions of increments X_j^i on the common random effect:

$$L\left(\mathbf{X_n}\right) = \prod_{i=1}^{n} f\left(X_1^i,\, X_2^i, ..., X_{k_i}^i\right) =$$

$$= \prod_{i=1}^{n} \int_0^{\infty} \left[\prod_{j=1}^{k_i} f_{Norm}\left(X_j^i; \omega\nu_{x^i}(t;\gamma,\beta), \sigma\right)\right] f_{trunc}\left(\omega;\delta,\alpha\right) d\omega \tag{12}$$

3 The analysis of the ILPN-134 lasers degradation data

We have considered the construction of the Wiener degradation model for estimating the reliability of the semiconducting laser module ILNP-134. The experiment data were presented by Zhuravleva, Ivanov and others in [9] and were described in Section 1 of the current paper. The detailed data analysis showed that two types of failure were observed during the accelerated test: the first type is related to the aging of the laser diode (LD) and the second one is related to the misalignment of the optical system (OS). Concerning to this, we decided to use the covariate function with the regression parameter β dependent on the covariate values.

At first, we have selected the most appropriate trend and covariate functions for the lasers data using the "fixed-effect" Wiener degradation model. Two types of the loglinear covariate functions depending on the β definition were chosen for the investigation:

$$r_1\left(x;\beta\right) = \begin{cases} e^{\beta_1^1 + \beta_2 \frac{1000}{x+273.2}}, & x \le 85; \\ e^{\beta_1^2 + \beta_2 \frac{1000}{x+273.2}}, & x > 85. \end{cases} \tag{13}$$

$$r_2\left(x;\beta\right) = \begin{cases} e^{\beta_1 + \beta_2^1 \frac{1000}{x+273.2}}, & x \le 85; \\ e^{\beta_1 + \beta_2^2 \frac{1000}{x+273.2}}, & x > 85. \end{cases} \tag{14}$$

Additionally, the exponential and power trend functions were considered:

$$\nu_1\left(t;\gamma\right) = \gamma_0 + \gamma_1 e^{t/r(x;\beta)}; \tag{15}$$

$$\nu_2\left(t;\gamma\right) = \gamma_0 + \gamma_1^{\gamma_2 t/r(x;\beta)}. \tag{16}$$

The parameters estimates, AIC and BIC values obtained for the "fixed-effect" Wiener degradation model with the different combinations of the considered trend and covariate functions are presented in the Table 1.

Table 1: The parameter estimates of the "fixed-effect" Wiener degradation models, AIC and BIC values

Trend function	Covariate function	σ	γ_1	γ_2	β_1	β_2	AIC	BIC
(15)	(13)	1.70	1.97	-	β_1^1: 3.71 β_1^2: 2.54	-0.42	-597.85	-573.35
(15)	(14)	1.69	2.16	-	3.15	β_2^1: 0.33 β_2^2: 0.63	-597.83	-573.33
(16)	(13)	1.70	0.07	1.75	β_1^1: 0.93 β_1^2: -0.33	-0.22	-595.91	-566.52
(16)	(14)	1.69	0.15	1.81	0.81	β_2^1: 0.05 β_2^2: 0.45	-595.91	-566.52

As can be seen from Table 1, the AIC and BIC values are smaller for the cases with the exponential trend function (15). As to the covariate functions, the values of information criteria are very close to each other, however the model with covariate function (13) describes the data more accurately. So, we choose these functions for the model construction.

As the unit-to-unit variability can be observed in the investigated data, we have considered the Wiener degradation model with random effects, where the random parameter μ has the left truncated normal distribution with the shift parameter δ equal to 0. The estimation results are presented in Table 2, the parameter estimates of the corresponding "fixed-effect" model are given for the comparison.

Table 2: The parameter estimates of the "fixed-effect" and "random-effect" Wiener degradation models, AIC and BIC values

Wiener degradation model	σ	α	β_1	β_2	AIC	BIC
"Fixed-effect" model	1.70	-	3.71 2.54	-0.42	-597.85	-573.35
"Random-effect" model	2.09	3.89e-03	2.03 1.57	-0.16	-707.54	-687.95

Basing on the AIC and BIC values presented in Table 2, it can be concluded that the "random-effect" Wiener degradation model is more appropriate model for the reliabiity analysis of the laser module ILPN-134 data.

The trend functions corresponding to the constructed "random-effect" Wiener degradation model are demonstrated on Figures 4-6.

Conclusions

In this paper, we have considered the problems of constructing the Wiener degradation model for the analysis of the ILPN-134 lasers data. An interesting feature of

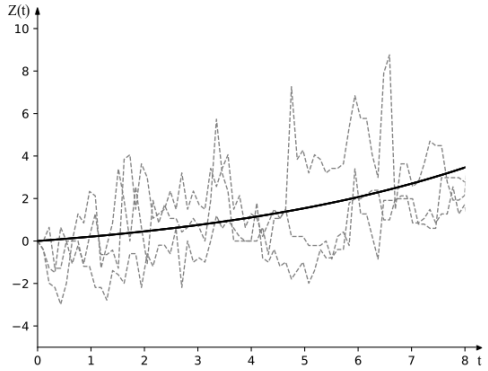


Figure 4: The trend function and observed degradation paths under 70°C

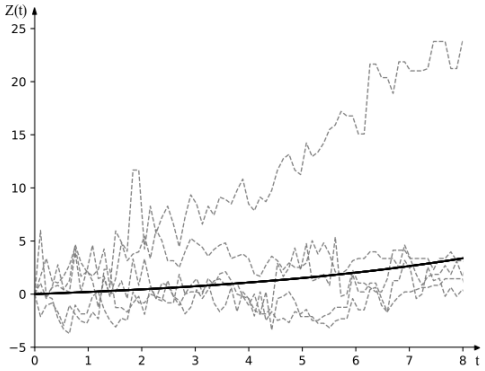


Figure 5: The trend function and observed degradation paths under 80°C

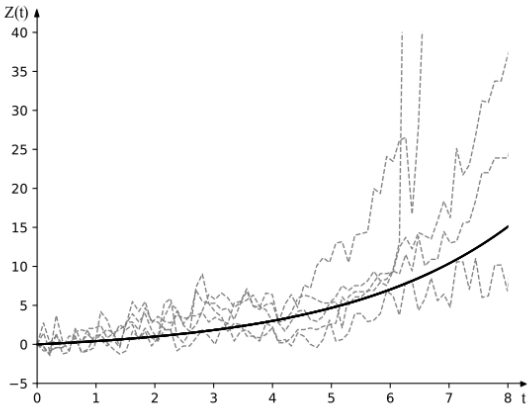


Figure 6: The trend function and observed degradation paths under 90°C

these data is that there are two types of failure, one of which arises only under high values of temperature. By this reason, we have taken the covariate function with the regression parameter dependent on the value of covarite (temperature). The exponential trend function was selected as the most appropriate one. Moreover, there is a significant unit-to-unit variability in data, and the Wiener degradation model with random effects turned out to be more appropriate than the corresponding "fixed-effect" model. The constructed model can be used for estimation of the reliability indicators, such as the probability of the non-failure operation during some period of time.

References

[1] Chetvertakova E., Chimitova E. (2018) *The Wiener degradation model with random effects in reliability metrology* Advanced Mathematical and Computational

Tools in Metrology and Testing XI. – Glasgow : World Scientific. pp. 162 - 169

[2] Lawless J., Crowder M. (2004) *Covariates and random effects on a gamma process model with application to degradation and failure*, Lifetime Data Analysis. Vol. 10. pp. 213-227

[3] Liao C.M., Tseng T. (2006) *Optimal design for step-stress accelerated degradation test*, IEEE Trans. Reliab. Vol. 55. pp. 59-66

[4] Meeker W.Q., L.A. Escobar, V. Chan (2002) *Using accelerated tests to preict service life in higly variable environments*, Service Life Prediction Methodolohies and Metrologies. pp. 16

[5] Nikulin M., Bagdonavicius N. (2001) *Accelerated Life Models: Modeling and Statistical Analisys*, Boca Raton : Chapman & Hall/CR. pp. 334

[6] Tsai C.-C., Tseng S.-T., Balakrishnan N. (2012) *Optimal Design for Degradation Tests Based on Gamma Processes with Random Effect*, IEEE Trans. Reliab. Vol. 61. pp. 604-613

[7] Wang X. (2010) *Wiener processes with random effects for degradation data*, Journal of Multivariate Analysis, Vol. 101, Is. 2, February 2010. pp. 340-351

[8] Whitmore G.A., F.Schenkelberg (1997) *Modeling Accelerated Degradation Data Using Wiener Diffusion with A Time Scale Transformation*, Lifetime Data Analysis. Vol. 3. pp. 27-45

[9] Zhuravleva O.V., Ivanov A.V., Kurnosov V.D., Kurnosov K.V., Romantsevich V.I., Chernov R.V. (2010) *Reliability estimation for semiconductor laser module ILPN-134*, Fizika i Tekhnika Poluprovodnikov, Vol. 44, No. 3. pp. 377–382.

Features of testing statistical hypotheses under Big Data analysis

B. Yu. Lemeshko, S. B. Lemeshko and M. A. Semenova
Novosibirsk State Technical University, Novosibirsk, Russia
e-mail: lemeshko@ami.nstu.ru

Abstract

The methods of construction of estimates are considered in the analysis of Big Data. The influence on the results of conclusions according to the Pearson Chi-squared test of choosing the number of intervals and grouping method is demonstrated. It is shows how the limited accuracy of data in large samples effects on the distribution of statistics of non-parametric tests. Recommendations on the application of tests under large samples analysis are given. It is shown that the distribution of statistics of tests for testing laws homogeneity, as well as the tests of homogeneity of the means and tests of homogeneity of the variances, is affected by the non-equilibrium character of the data presented in the compared samples.

Keywords: Big Data; parameter estimation; testing hypotheses; goodness-of-fit tests; homogeneity tests; statistical simulation

Introduction

The questions of application of statistical methods to the analysis of large data arrays (Big Data) are of great interest in recent years. In connection with the rapid accumulation of gigantic volumes of information, there is a need for research the accumulated data, for finding, extracting and using the laws hidden in data, including probabilistic ones. Naturally, one can try to apply methods and tests from the vast arsenal of classical mathematical statistics for the analysis of big data, using popular software systems for statistical analysis. However, application of the classical apparatus of applied mathematical statistics for the analysis of big data, as a rule, leads to specific problems that limit the possibilities of correct application of this apparatus.

In this paper, we will discuss only the methods and tests associated with the analysis of one-dimensional random variables, the real problems of which are most familiar to us. At least three situations can be considered where increasing sample size causes problems in application of methods or tests.

Firstly, due to the "curse of dimension", well-proven methods and algorithms become ineffective. In particular, problems arise under the calculation of estimates of parameters. When using estimation methods that operate on non-grouped data, the computational costs increase cardinally with increasing size of samples analyzed. The convergence of iterative algorithms used in estimation worsens. A significant factor is no robustness of certain types of estimation. The natural way to resolve this situation is the use of estimation methods that involve grouping data [1]. But in this case, the

question arises: how the estimates obtained for grouped data will affect the properties of hypotheses tests in which estimates will be used. For example, how will this affect the statistics distributions of non-parametric goodness-of-fit tests when testing composite hypotheses? In this case, the statistic distributions significantly depend on the method of parameter estimation [2, 3, 4, 5].

Secondly, a lot of popular statistical tests are not adapted even for samples of about thousand observations, since the information on the distributions of statistics of these tests is presented only by brief tables of critical values for some sample sizes n. By rough estimate, the count of such tests is more than 80% of all tests count. It should be noted that the possibility of application such tests with "reasonable" values of sample size is easily resolved by statistical simulation of distributions of statistics for given sample size and validity of the tested hypothesis H_0. This simulation can be carried out interactively during statistical analysis [6, 7]. The empirical distribution $G_N(S_n\,|H_0)$ of statistic S of test constructed as a result of simulation with size N can then be used to estimate the achieved significance level p_{value} by the value of the statistics S^* calculated from the analyzed sample.

Thirdly, the application of tests, for which the limiting (asymptotic) distributions of statistics are known, always leads to rejection of even true tested hypothesis with increasing sample sizes. This is typical, for example, for goodness-of-fit tests, for a lot of special tests for testing hypotheses of normal distribution, uniform distribution or exponential distribution, etc. In [8], it has been shown that this problem is associated not only and not so much with the increasing computational costs, as with the limited accuracy of the analyzed data (with limited measurement accuracy). A similar problem hinders the correctness of application of homogeneity tests (homogeneity of laws, homogeneity of variance, to a lesser degree of homogeneity of means) under large samples. As will be shown, in the case of homogeneity tests, the reason lies in the unevenness of measurements in the analyzed samples.

1 Estimation of the parameters of distribution

Estimates of the parameters of distributions can be obtained by various methods. The maximum likelihood estimates (MLE) characterized by the best asymptotic properties and calculated by maximizing the likelihood function

$$\hat{\theta} = \arg\max_{\theta} \prod_{j=1}^{n} f(x_j, \theta), \tag{1}$$

or by maximizing the logarithm of this function, where θ is unknown parameter (generally vector), $f(x, \theta)$ is the density function of the distribution law, $x_1, x_2, ..., x_n$ are sample observation. For some laws, the distribution of MLE of parameters is obtained as statistics simply computed from the observations of the samples, but in most cases MLE are the result of using some iterative method.

When calculating MD-estimates (estimates of the minimum distance), some measure of proximity (distance) $\rho(F(x,\ \theta),\ F_n(x))$ between the theoretical $F(x,\ \theta)$ and

empirical $F_n(x)$ distributions is minimized. MD-estimates can be obtain as a result of solving following task

$$\hat{\theta} = \arg \min_{\theta} \rho(F(x,\theta), F_n(x)). \tag{2}$$

For example, the statistics of nonparametric goodness-of-fit tests (Kolmogorov, Cramer-von Mises-Smirnov, Anderson-Darling, Kuiper, Watson, and others [9]) can be used as measures of proximity.

With relatively small sample sizes, L-estimates of parameters can be used. These estimates are some linear combinations of order statistics (elements of variational series $X_{(1)} < X_{(2)} < \ ... < \ X_{(n)}$ constructed from original sample $x_1, x_2, ..., x_n$).

MLE of parameters of distribution, as a rule, are not robust. The presence of anomalies of sample observations or the inaccuracy of the assumption about the form of distribution leads to the construction of models with distribution functions that are unacceptably deviating from empirical distributions. MD-estimations have greater stability.

Obviously, the calculation of estimates (1) and (2) is associated with serious computational difficulties for very large samples. In the case of grouped sample, the sample observations are associated with a set of non-intersecting intervals, which divide the domain of definition of a random variable into k non-intersecting intervals by boundary points

$$x_{(0)} < x_{(1)} < \ldots < x_{(k-1)} < x_{(k)},$$

where $x_{(0)}$ is the lower bound of the domain of definition of random variable X; $x_{(k)}$ is the upper bound of the domain of definition of random variable X.

MLE by grouped sample [1] are calculated by maximizing the likelihood function

$$\hat{\theta} = \arg\max_{\theta} \prod_{i=1}^{k} P^{n_i}(\theta), \tag{3}$$

(3) where $P_i(\theta) = \int\limits_{x_{(i-1)}}^{x_{(i)}} f(x,\ \theta)dx$ is the probability of the observation entering in the i-th interval of values, n_i is the number of observations that fell into the i-th interval, $\sum\limits_{i=1}^{k} n_i = n$. Estimates by grouped samples can be obtained by minimizing statistics χ^2

$$\hat{\theta} = \arg\min_{\theta}\ n \sum_{i=1}^{k} \frac{(n_i/n - P_i(\theta))^2}{P_i(\theta)}, \tag{4}$$

as well as other statistics. In [10], it was shown that all of estimation method for grouped data considered give consistent and asymptotically effective estimates under appropriate regularity conditions. However, the most preferred estimates are MLE. An important advantage of estimates based on grouped data is robustness [11].

In the case of presence of non-grouped data, estimates for grouped data are rarely applied. This is due to the greater computational costs and necessity to numerical integration in the computation $P_i(\theta)$, that requires appropriate software support.

In the case of large sample sizes, the situation changes. Computational costs do not change as computations grow with a fixed number of grouping intervals, but increase only with an increase in the number of intervals k. This means that it is advisable to use MLE by grouped samples in the conditions of Big Data. These are robust and asymptotically efficient estimates. The quality of estimates for small k can be improved using asymptotically optimal grouping (AOG) [1, 12, 13], in which the losses in Fisher information associated with grouping are minimized.

2 Application of χ^2-test under large samples

The statistic of Pearson χ^2 goodness-of-fit test has the following form

$$X_n^2 = n \sum_{i=1}^{k} \frac{(n_i/n - P_i(\theta))^2}{P_i(\theta)}. \tag{5}$$

In the case of testing simple hypothesis, this statistic obeys χ_r^2-distribution with $r = k - 1$ degrees of freedom if $n \to \infty$ and the null hypothesis is true.

In the case of testing composite hypothesis and estimating m parameters of distribution by sample statistic (4) obeys χ_r^2-distribution with $r = k - m - 1$ degrees of freedom, if the estimates are obtained by minimizing (4) these statistics, or using MLE (3) (or other asymptotically effective estimates for grouped data).

The distribution of statistic (5) does not agree with χ_{k-m-1}^2–distribution when parameter estimations are obtained by non-grouped data. It is recommended to apply the Nikulin-Rao-Robson test when MLE were obtained according to ungrouped data [14, 15].

There are not principal problems that prevent application of Pearson χ^2-test under Big Data. Only computational difficulties are possible.

Let us illustrate the results of application Pearson χ^2-test on the example of testing hypothesis of normal distribution with density

$$f(x, \theta) = \frac{1}{\theta_1 \sqrt{2\pi}} \exp\left\{ -\frac{(x - \theta_0)^2}{2\theta_1^2} \right\}.$$

by sufficiently large sample. The sample of $n = 10^7$ observations was modeled according to the standard normal law $N(0, 1)$ ($\theta_0 = 0$, $\theta_1 = 1$).

In Table 1, there are the results of testing simple hypotheses about standard normal law $N(0, 1)$ with various numbers of intervals in the case of equal-frequency grouping (EFG) and $k = 15$ in the case of asymptotically optimal grouping (AOG).

In the case of AOG, the power of Pearson χ^2-test maximizes for close competing laws [16, 17, 18]. The table shows the values X_n^{2*} of statistics (5), which calculated by the sample, and the corresponding values $p_{value} = P\{X_n^2 \geq X_n^{2*} | H_0\}$ of the achieved significance level. As you can see, the results depend on both the splitting method and the number of intervals. The power of test also depends on these factors [19].

Table 1: Results of testing simple hypothesis about $N(0,1)$

	AOG	EFG						
k	15	15	50	75	100	500	1000	2000
X_n^{2*}	7.75162	9.18380	56.8942	79.4904	96.5701	493.995	1044.57	2099.91
p_{value}	0.90186	0.81910	0.20475	0.31026	0.55038	0.55482	0.15403	0.05702

Table 2 shows the results of testing composite hypotheses. MLE $\hat{\theta}_0$ and $\hat{\theta}_1$ obtained for grouped data with the corresponding number of intervals k, statistics values X_n^{2*} and p_{value} are presented.

MLE of parameters by complete ungrouped sample are $\hat{\theta}_0 = 0.000274$ and $\hat{\theta}_1 = 1.000177$. In [20, 21], models of distributions of statistic (5) were constructed for the case of testing composite hypothesis of normal law using MLE by ungrouped data and AOG. The value of statistic calculated by the sample is $X_n^{2*} = 6.600521$ for $k = 15$, the estimate of p-value obtained in accordance with the limit distribution model given in [20, 21] is $p_{value} = 0.886707$. These values indicate a good agreement between the complete sample and the normal law $N(0.000274, 1.000177)$.

Table 2: Results of testing composite hypothesis

	AOG	EFG						
k	15	15	50	75	100	500	1000	2000
$\hat{\theta}_0$	0.00028	0.00030	0.000244	0.00027	0.00027	0.00028	0.00027	0.00027
$\hat{\theta}_1$	1.00715	1.00263	1.00173	1.00134	1.00112	1.00039	1.00031	1.00024
X_n^{2*}	927.920	99.9963	101.767	104.511	112.151	493.716	1043.47	2098.61
p_{value}	0.0	5.58e-16	6.50e-06	0.00739	0.13938	0.53317	0.14922	0.05572

It should be noted that the MLE by grouped sample for $k = 2000$ and the MLE by ungrouped sample are very close. At the same time, p-value for $k = 2000$ is much lower than 0.886707.

Thus, the result of testing composite hypotheses using Pearson χ^2-test significantly depends on the number of intervals k.

3 Nonparametric goodness-of-fit tests under big samples

If one can omit the growth of computational difficulties, the main reason for possible non-correctness of conclusions by big data using non-parametric goodness-of-fit tests is the limited accuracy of the data in large sample.

As a rule, volumes of samples in Big Data (belonging to some continuous distribution law) are practically unlimited, but the observations itself are presented with

limited accuracy (rounded with some Δ). In essence, there is "violation of assumption" that a continuous random variable is observed.

Suppose, the goodness-of-fit test with statistic S is used to test a simple hypothesis $H_0: \ F_n(x) = F(x)$, where $F_n(x)$ is empirical distribution constructed from sample

$$x_1, x_2, ..., x_n$$

of n observations. Suppose, there is limit distribution of statistic $G(S\,|H_0)$ for this goodness-of-fit test. In the case of trueness of H_0, the empirical distribution $F_n(x)$ corresponding to sample of continuous random variables (without rounding) converges to the distribution function of this random variable $F(x)$ for $n \to \infty$. The empirical distribution of statistics $G_N(S_n\,|H_0)$ based on samples of continuous random variable for $n \to \infty$ (and the number of simulation experiments $N \to \infty$) converges to the limit distribution $G(S\,|H_0)$ of this statistics.

However, the measurement results are rounded off (fixed) with some Δ. Therefore, $\max|F_n(x) - F(x)|$ will cease to decrease starting with certain n, depending on $F(x)$, domain of definition of the random variable and Δ. The distribution $G_N(S_n\,|H_0)$ will deviate from the limiting distribution $G(S\,|H_0)$ with increasing n (the more Δ, that the less n).

The results of studies for demonstrating the effect of accuracy of data on the distribution of statistics will be shown on 3 classical goodness-of-fit tests.

The Kolmogorov test statistics is used with the Bolshev correction[9]

$$S_K = \sqrt{n}D_n + \frac{1}{6\sqrt{n}} = \frac{6nD_n + 1}{6\sqrt{n}}, \tag{6}$$

where $D_n = \max(D_n^+,\ D_n^-)$, $D_n^+ = \max\limits_{1\le i\le n}\left\{\frac{i}{n} - F(x_i,\ \theta)\right\}$,
$D_n^- = \max\limits_{1\le i\le n}\left\{F(x_i,\ \theta) - \frac{i-1}{n}\right\}$; n is the number of observations; $x_1,\ x_2, \ldots,\ x_n$ are sample values ordered ascending; $F(x,\ \theta)$ is distribution function of law tested. The distribution of S_K under simple hypothesis in the limit obeys the Kolmogorov law with the distribution function $K(S)$ [9].

The Cramer-von Mises-Smirnov test statistic is

$$S_\omega = \frac{1}{12n} + \sum_{i=1}^{n}\left\{F(x_i,\ \theta) - \frac{2i-1}{2n}\right\}^2 \tag{7}$$

and under testing simple hypothesis this statistic allows to law with distribution function $a1(s)$ [9]. The Anderson-Darling test statistic has the following form [22]

$$S_\Omega = -n - 2\sum_{i=1}^{n}\left\{\frac{2i-1}{2n}\ln F(x_i,\ \theta) + \left(1 - \frac{2i-1}{2n}\right)\ln(1 - F(x_i,\ \theta))\right\}. \tag{8}$$

In the case of testing simple hypothesis this statistic allows to law with distribution function $a2(s)$ [9].

In [8], the distributions of statistics (6)-(8) of nonparametric goodness-of-fit tests were studied depending on the accuracy of recording the observed values of random

variables. The number of significant decimal places, to which the observed values were rounded, was set. This determined the number of unique values that could be in the generated samples. As a rule, the number of simulation experiments carried out to simulate the empirical distributions of statistics was $N = 10^6$.

The deviation of real (empirical) distribution of statistics from the limit distribution was studied by evaluating median $\tilde{S}_n$ of empirical distribution of statistics obtained as a result of modeling. If real distribution of statistics with sample sizes n does not deviate from the limit distribution, then the probability $P\left\{S > \tilde{S}_n\right\}$ calculated from the corresponding limit distribution is 0.5. If real distribution of statistics shifts to large area of values (to the right of the limit distribution), the estimates $\hat{p}_v = P\left\{S > \tilde{S}_n\right\}$ are decrease. One can judge the correctness of achieved significance level p_{value} calculated from the limit distribution of statistics (in the case of testing simple hypotheses, respectively, by $K(S)$, $a1(S)$ and $a2(S)$) by the value of deviation of estimates $\hat{p}_v$ from 0.5.

When rounding to within 1 in samples belonging to $N(0,1)$, 9 unique values may appear, when rounding to within $\Delta = 0.1$ about 86 unique values, with accuracy $\Delta = 0.01$ – about 956, to within $\Delta = 0.001$ – about 9830.

As the simulation results showed [8], when rounding up observations to integer values, the use of limit distributions of test statistics is **absolutely** excluded.

The distributions of statistic of Kolmogorov test $G(S_n\,|H_0)$ is essentially discrete under $\Delta = 0.1$. The deviation $G(S_n\,|H_0)$ from the limit distribution $K(S)$ for $\Delta = 0.1$ should be taken into account already for $n > 20$, $\Delta = 0.01$ – for $n > 250$, and if $\Delta = 0.001$ the value $n_{\max}$ shifts to value about 10^4. In the case of Cramer-von Mises-Smirnov and Anderson-Darling tests, the deviation $G(S_n\,|H_0)$ from the limit $a1(S)$ and $a2(S)$ for $\Delta = 0.1$ should be taken into account for $n > 30$, $\Delta = 0.01$ – for $n > 1000$, and if $\Delta = 0.001$ – the value $n_{\max}$ shifts to 5×10^5.

Figure 1 shows the dependence of distributions of statistics (7) of Cramer-von Mises-Smirnov test on the degree of rounding Δ at sample size $n = 1000$ for the case of testing simple hypothesis about standard normal law. The limit distribution $a1(S)$, that occurs in the case without rounding, as well as real distributions of statistics $G(S_{1000}\,|H_0)$ at degree of rounding $\Delta = 0.01,\ 0.05,\ 0.1,\ 0.2,\ 0.3$. As you can see if $\Delta = 0.01$ distribution $G(S_{1000}\,|H_0)$ does not practically differ from $a1(S)$, but with increasing Δ deviation $G(S_{1000}\,|H_0)$ from $a1(S)$ rapidly increases.

Consequently, in order to analyze large samples using the appropriate nonparametric goodness-of-fit tests with corresponding limit distributions, statistics should be calculated not over the sample, but according to samples extracted by uniform law from general population (original sample analyzed). The size of extracted sample should take into account the accuracy of the data (the number of possible unique values in the sample) and not exceed certain value $n_{\max}$ at which (for given accuracy) the distribution of test statistics $G(S_{n_{\max}}\,|H_0)$ does not really differ from the limit distribution $G(S\,|H_0)$.

In the case of testing composite hypotheses, the tested hypothesis has the form $H_0 : F(x) \in \{F(x,\,\theta), \theta \in \Theta\}$, where Θ is domain of parameter θ definition. If

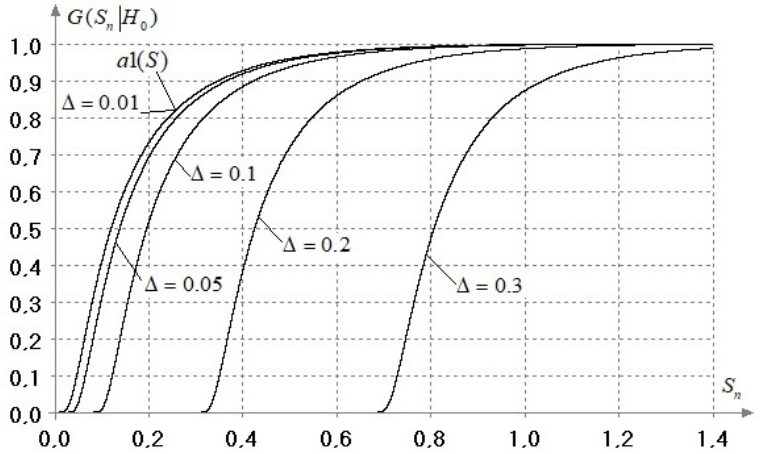


Figure 1: Statistic distributions $G(S_n\,|H_0)$ of Cramer-von Mises-Smirnov test depending on Δ for $n = 1000$

the estimate $\hat{\theta}$ of scalar or vector parameter of law is based on the same sample that the hypothesis is tested on, then the distribution of statistics $G(S\,|H_0)$ for any nonparametric goodness-of-fit test differs significantly from the limit distribution for testing simple hypothesis [23]. If estimates of parameters obtain by the same sample that hypothesis tested, the following factors influence the distribution of statistics $G(S\,|H_0)$ [24]: distribution law $F(x,\,\theta)$ corresponding to the true hypothesis H_0; type of estimated parameter and the number of estimated parameters; in some situations, specific values of parameter (for example, in the case of gamma distribution, etc.); used parameter estimation method.

Obviously, in the case of testing composite hypotheses, we encounter the same problems and must extract sample of size $n < n_{\max}$ from "general population" in order to use when analyzing Big Data with limited accuracy of fixed data. For example, it should be do for application of models of limit distributions of test statistics when testing composite hypotheses [2, 3, 4, 5, 24].

It should be noted, if the estimation $\hat{\theta}$ of parameter is found by one of the above methods by the entire big data array, and then the test is applied to the sample of size $n < n_{\max}$ extracted from the same array, then when testing hypothesis H_0 : $F(x) = F(x, \hat{\theta})$, where $\hat{\theta}$ is previously obtained estimate, the distribution of statistics $G(S\,|H_0)$ will as in the case of testing simple hypothesis.

All of the above fully applies to application of nonparametric Kuiper [25] and Watson [26, 27] goodness-of-fit tests by big samples. The distributions of statistics of third Zhang goodness-of-fit tests [28], which are based on Kolmogorov, Cramer-von Mises-Smirnov and Anderson-Darling tests, depend on sample sizes n. Therefore, there can be no talk about application of limit distributions of statistics. However, distribution of statistics $G(S_n\,|H_0)$ in the same way depends on degree of rounding Δ. Consequently, the critical values of statistics obtained for continuous random variables and n cannot be used with the same n, but with significant degree of rounding Δ. The

problem can be resolved by statistical modeling (including, in the interactive mode [6, 7]) of statistical distributions for given n and Δ with the trueness of the tested hypothesis H_0. The empirical distribution of $G_N(S_n\,|H_0\,)$ statistics S of corresponding test constructed as a result of N simulation experiments under these conditions can be used to estimate the achieved significance level p_{value}. That is how this problem is solved in the ISW software system being developed [29].

4 Other goodness-of-fit tests under big samples

It should be noted that the degree of rounding of recorded data affects properties of other tests in similar way. In particular, special tests aimed for testing the hypothesis about normal law, uniform law, or exponential law, etc.

It should be noted that in the conditions of large samples (in the presence of repeated observations), a lot of good tests turn out to be inoperable. This is due to the fact that the type of statistics of these tests excludes the presence of repeated observations (or the number of repeated values greater than the size of the "m window" used in statistics). This note concerns tests using entropy estimates (Vacicek [30] and Alizadeh Noughabi [31] normality tests, Dudewics-van der Meulen [32] and Zamanzade [33] uniformity tests), as well as new goodness-of-fit tests using estimates of Kullback-Leibler information [34].

5 Homogeneity tests under big samples

In the case of multi-sample tests, which include homogeneity tests, 2 or more samples are compared. The distributions of statistics of multi-sample tests are influenced by non-uniformity of data presented in the analyzed samples. The two-sample Lehmann-Rosenblatt homogeneity test was proposed in [35] and studied in [36]. Statistic based on two samples $x_{11}, x_{12}, ..., x_{1,n_1}$ and $x_{21}, x_{22}, ..., x_{2,n_2}$:

$$S_{LR} = \frac{1}{n_1 n_2 (n_1 + n_2)} \left[n_1 \sum_{i=1}^{n_1} (r_i - i)^2 + n_2 \sum_{j=1}^{n_2} (s_j - j)^2 \right] - \frac{4 n_1 n_2 - 1}{6(n_1 + n_2)}, \qquad (9)$$

where r_i is serial number (rank) of x_{1i}; s_j is serial number (rank) of x_{2j} in the united variation range.

The limit distribution of statistic (9) under true tested hypothesis $H_0: \ F_1(x) = F_2(x)$ is the same distribution $a1(s)$ [36], which is limit for statistic of Cramer-von Mises-Smirnov goodness-of-fit test.

Let us consider how degree of rounding affects distribution of statistic of homogeneity tests in the case of true H_0 and belonging of analyzed sample observations to the standard normal law.

Figure 2 demonstrates the dependence of distribution of statistic $G(S_{LR}\,|H_0)$ of Lehmann-Rosenblatt homogeneity test on degree of rounding Δ_2 of observations in the second sample when rounding in the first sample $\Delta_1 = 0.01$. The sample sizes are $n_i = 1000$.

The deviation $G(S_{LR}\,|H_0)$ from $a1(S)$ turns out to be significant already for $\Delta_2 = 0.05$. The deviation $G(S_{LR}\,|H_0)$ from $a1(S)$ rapidly increases with increasing sample sizes for fixed Δ_2. The deviation increases with Δ_2 growth for fixed sample size. The distributions of statistic $G(S_{LR}\,|H_0)$ of Lehmann-Rosenblatt homogeneity test depend on the difference between Δ_1 and Δ_2.

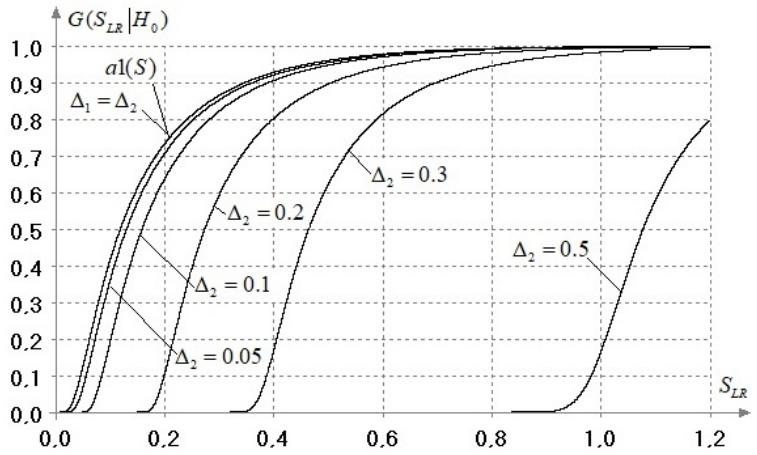


Figure 2: Statistic distributions of Lehmann-Rosenblatt homogeneity test depending on Δ_2 for $\Delta_1 = 0.01$ and $n_i = 1000$

Similarly, the distributions of other two-sample homogeneity tests (Smirnov, Anderson-Darling-Pettitt) depend on the difference between Δ_1 and Δ_2. It is natural that the distributions of statistics of all multi-sample tests of homogeneity (set of which is considered in [37]) depend on the non-equivalence of data presentation in the analyzed samples.

The distributions of statistic of parametric tests of homogeneity of means do not suffer from such dependence on degree of rounding of measurements as tests of homogeneity of laws considered above. At the same time, it should be noted that the power of tests decreases with decrease of accuracy of data recorded.

The distributions of statistic of parametric tests of homogeneity of variances, unlike tests of homogeneity of means, are more dependent on degree of rounding. In some ways, this is due to the greater sensitivity of the variance estimates to the accuracy of measurement results.

Parametric tests of Cochran, Bartlett, Fisher, Hartley, Neumann-Pearson and Overall-Woodward Z-test are the most preferable in terms of power among the set of parametric and non-parametric tests of homogeneity of variances. These tests are equivalent in power in the case of two sample and fulfilling the assumption that analyzed samples are normal. But in the case of k sample, the power advantage turns out to be Cochran test has power advantage [38, 39, 40, 41]. Statistic of Cochran test [42] can be written as

$$Q = \frac{S^2_{\max}}{S_1^2 + S_2^2 + \cdots + S_k^2},$$

where $S^2_{\max} = \max(S^2_1,\ S^2_2, \ldots,\ S^2_k)$; k is the number of samples; $S^2_i,\ \ i = \overline{1,k}$, are the estimates of variances obtained by samples. Tested hypothesis $H_0:\ \sigma^2_1 = \sigma^2_2 = \ldots = \sigma^2_k$ deviates for large values of statistic. The distributions of statistic $G(Q_n\,|H_0)$ of Cochran test depend on the number of compared samples k and the sizes of these samples n_i.

Figure 3 illustrates the dependence of the distribution of statistics $G(Q_n\,|H_0)$ of Cochran test on degree of rounding of observations in the second sample Δ_2 without rounding in the first sample ($\Delta_1 = 0$). Sample sizes are $n_i = 1000$ and k=2. As can be seen, the dependence of the distribution $G(Q_n\,|H_0)$ on large (different) degrees of rounding Δ_1 and Δ_2 is very significant.

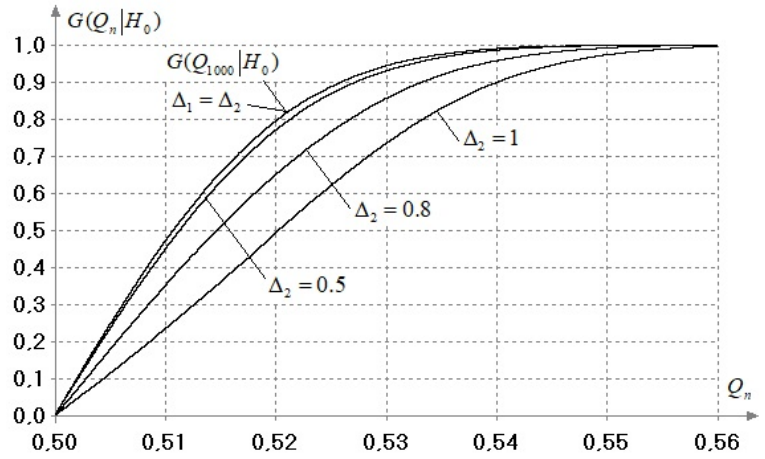


Figure 3: Statistic distributions $G(Q_n\,|H_0)$ of Cochran homogeneity test depending on Δ_2 for $\Delta_1 = 0$ and $n_i = 1000$

The limited accuracy of measurements always leads to decrease of the power of homogeneity tests. The drop in the power of Cochran test with increasing degree of rounding (with equal Δ_i , equal sample sizes $n_1 = n_2 = 100$, and k=2) is shown in Table 3. The competing hypothesis has the form $H_1:\ \ \sigma_2 = 1.2\sigma_1$. Also this table shows power of Klotz nonparametric test [43]. It is interesting that with increasing Δ_i the power of nonparametric test decreases faster than power of parametric one.

Let us emphasize that, similarly, the value of rounding Δ_i affects the distributions of statistics and the power of other tests of homogeneity of variances.

So, the distributions of statistics $G(S\,|H_0)$ of parametric tests of homogeneity of variances with the same degree of rounding Δ_i of measurement in the analyzed samples do not differ from corresponding distributions without rounding ($\Delta_i = 0,\ \ i = \overline{1,k}$). However, the same distributions with different Δ_i differ significantly from distributions without rounding.

In the case of trueness of competing hypotheses, degree of rounding Δ_i (measurement registration accuracy) has significant impact on the distributions of statistics and on the power relative to these competing hypotheses (including under equal Δ_i

Table 3: Estimates of power of Cochran and Klotz tests under H_1

	Cochran test			
α	Without rounding	$\Delta_1 = \Delta_2 = 0.1$	$\Delta_1 = \Delta_2 = 0.2$	$\Delta_1 = \Delta_2 = 0.5$
0.1	0.564	0.562	0.560	0.550
0.05	0.438	0.435	0.434	0.424
	Klotz test			
α	Without rounding	$\Delta_1 = \Delta_2 = 0.1$	$\Delta_1 = \Delta_2 = 0.2$	$\Delta_1 = \Delta_2 = 0.5$
0.1	0.540	0.539	0.535	0.504
0.05	0.413	0.412	0.407	0.378

in samples). Similar conclusions hold for the entire set of parametric tests of homogeneity of variances considered in [37].

Conclusions

It is advisable to use parameter estimation methods involving the grouping of data for constructing probabilistic models by big samples. Such estimates are robust, and computational costs do not depend on sample sizes in contrast to estimates by ungrouped data.

There are no serious objections to application of Pearson χ^2-test for analysis of big samples. This test retains both its positive qualities and its inherent flaws.

The main problem preventing the correct application of nonparametric goodness-of-fit tests for analysis of big samples is limited accuracy of data representation. Due to limited accuracy with increasing sample volumes, the real distributions of statistics deviate from the limit ones that occur under the assumption of continuity of observed random variables. Therefore, the application of classical results for corresponding tests may lead to incorrect conclusions. On the one hand, it is possible to recommend application of these tests to samples extracted from Big data, the size of these samples is limited by accuracy of presenting data analyzed (the number of possible unique values in the sample). On the other hand, it is possible to propose the use of statistical modeling methods to estimate real distributions of test statistics $G_N(S_n\,|H_0)$ (corresponding to degree of rounding Δ of data in sample analyzed) and then use $G_N(S_n\,|H_0)$ to estimate achieved significance level p_{value}.

The reason for possible incorrectness of conclusions when using classical results concerning the distributions of statistics of corresponding homogeneity tests may be the non-equilibrium measurement in the compared samples. Statistical modeling can be proposed to simulate actual distribution of statistics $G_N(S_n\,|H_0)$ of test applied (with appropriate degrees of rounding Δ_i and sizes n_i of compared samples). The distribution $G_N(S_n\,|H_0)$ obtained can then be used to estimate achieved significance level p_{value}.

Similar methodology of analysis of big samples is implemented in ISW software system [29].

Acknowledgements

The research is supported by the Russian Ministry of Education and Science (projects No 1.4574.2017/6.7 and No 1.1009.2017/4.6).

References

[1] Kulldorff G. Contributions to the theory of estimation from grouped and partially grouped samples. Almqvist&Wiksell. 1961.

[2] Lemeshko B.Yu, Lemeshko S.B. Distribution models for nonparametric tests for fit in verifying complicated hypotheses and maximum-likelihood estimators. P. I // Measurement Techniques. 2009. Vol. 52, No. 6. – P. 555-565.

[3] Lemeshko B.Yu., Lemeshko S.B. Models for statistical distributions in nonparametric fitting tests on composite hypotheses based on maximum-likelihood estimators. P. II // Measurement Techniques. 2009. Vol. 52, No. 8. – P. 799-812.

[4] Lemeshko B.Yu., Lemeshko S.B., Postovalov S.N. Statistic Distribution Models for Some Nonparametric Goodness-of-Fit Tests in Testing Composite Hypotheses // Communications in Statistics – Theory and Methods. 2010. Vol. 39, No. 3. – P. 460-471.

[5] Lemeshko B.Yu., Lemeshko S.B. Construction of Statistic Distribution Models for Nonparametric Goodness-of-Fit Tests in Testing Composite Hypotheses: The Computer Approach // Quality Technology & Quantitative Management. 2011. Vol. 8, No. 4. – P. 359-373.

[6] Lemeshko B.Yu., Lemeshko S.B., Rogozhnikov A.P. Real-Time Studying of Statistic Distributions of Non-Parametric Goodness-of-Fit Tests when Testing Complex Hypotheses // Proceedings of the International Workshop "Applied Methods of Statistical Analysis. Simulations and Statistical Inference" – AMSA'2011, Novosibirsk, Russia, 20-22 September, 2011. – P. 19-27.

[7] Lemeshko B.Yu., Lemeshko S.B., Rogozhnikov A.P. Interactive investigation of statistical regularities in testing composite hypotheses of goodness of fit // Statistical Models and Methods for Reliability and Survival Analysis : monograph. – Wiley-ISTE, 2013. – Chap. 5. – P. 61-76.

[8] Lemeshko B.Yu., Lemeshko S.B., Semenova M.A. To Question of the Statistical Analysis of Big Data // Tomsk State University Journal of Control and Computer Science. 2018. 44. – P. 40-49. DOI: 10.17223/19988605/44/5

[9] Bolshev L.N., Smirnov N.V. Tables for Mathematical Statistics. Moscow: Nauka. 1983. – 416 p.

[10] Rao S.R. Linear statistical methods and their applications. Moscow: Nauka. 1968. – 548 p.

[11] Lemeshko B.Yu. Grouping observations as a way to generate robust estimates // Nadezhnost' i kontrol' kachestva. 1997. 5. – P. 26-35.

[12] Denisov V.I., Lemeshko B.Yu., Tsoi E.B. Optimal grouping, parameter estimation, and planning regression experiments. In 2 parts // Novosibirsk : NSTU Publisher, 1993. – 347 p.

[13] Lemeshko B.Yu., Lemeshko S.B., Postovalov S.N., Chimitova E.D. Statistical Data Analysis, Simulation and Study of Probability Regularities. Computer Approach / B.Yu. Lemeshko, S.B. Lemeshko, S.N. Postovalov, E.V. Chimitova. Novosibirsk : NSTU Publisher, 2011. – 888 p.

[14] Nikulin M.S. On chi-square testing for continuous distributions // Theory of Probability & Its Application. 1973. Vol. 18. No. 3. – P. 675-676.

[15] Rao K.C., Robson D.S. A chi-squared statistic for goodness-of-fit tests within the exponential family // Commun. Statist. 1974. Vol. 3. – P. 1139-1153.

[16] Denisov V.I., Lemeshko B.Yu. Optimal grouping in the processing of experimental data // Measuring Information Systems. Novosibirsk. – P. 5-14.

[17] Lemeshko B.Yu. Asymptotically optimal grouping of observations is to ensure the maximum power of the tests // Nadezhnost' i kontrol' kachestva. 8. – P. 3-14.

[18] Lemeshko B.Yu. Asymptotically optimum grouping of observations in goodness-of-fit tests // Industrial laboratory. Diagnostics of materials. 1998. 64(1). – P. 56-64.

[19] Lemeshko B.Yu., Chimitova E.V. On the choice of the number of intervals in the goodness-of-fit tests of type χ^2 // Industrial laboratory. Diagnostics of materials. 2003. 69(1). – P. 61-67.

[20] Lemeshko B.Yu. Tests for checking the deviation from normal distribution law. Moscow: INFRA–M. 2015. DOI: 10.12737/6086

[21] Lemeshko B.Yu. Chi-Square-Type Tests for Verification of Normality // Measurement Techniques, 2015. Vol. 58, No. 6. – P. 581-591. DOI: 10.1007/s11018-015-0759-2

[22] Anderson T.W., Darling D.A. A test of goodness of fit // J. Amer. Statist. Assoc. 1954. Vol. 29. – P. 765-769.

[23] Kac M., Kiefer J., Wolfowitz J. On tests of normality and other J. tests of goodness of fit based on distance methods // Ann. Math. Stat. 1955. Vol. 26. – P. 189-211.

[24] Lemeshko B.Yu. Nonparametric goodness-of-fit tests. Guide on the application. M: INFRA–M, 2014. – 163 p. DOI: 10.12737/11873

[25] Kuiper N.H. Tests concerning random points on a circle // Proc. Konikl. Nederl. Akad. Van Wettenschappen. 1960. Series A. V.63. – P. 38-47.

[26] Watson G.S. Goodness-of-fit tests on a circle. I // Biometrika. 1961. V. 48. No. 1-2. – P. 109-114.

[27] Watson G.S. Goodness-of-fit tests on a circle. II // Biometrika. 1962. V. 49. No. 1-2. – P. 57- 63.

[28] Zhang J. Powerful goodness-of-fit tests based on the likelihood ratio // Journal of the Royal Statistical Society: Series B. 2002. V.64. No. 2. – P. 281-294.

[29] ISW – Program system of the statistical analysis of one-dimensional random variables. URL: https://ami.nstu.ru/ headrd/ISW.htm. (address date 12.05.2019).

[30] Vacicek O. A test for normality based on sample entropy // Journal of the Royal Statistical Society: Series B. 1976. V. 38, No. 1. – P. 54-59.

[31] Alizadeh Noughabi H. A new estimator of entropy and its application in testing normality // Journal of Statistical Computation and Simulation. 2010. V. 80. No. 10. – P. 1151-1162.

[32] Dudewics E. J., van der Meulen E. C. Entropy-based test of uniformity // J. Amer. Statist. Assoc. 1981. V. 76. No. 376. – P. 967-974.

[33] Zamanzade E. Testing uniformity based on new entropy estimators // Journal of Statistical Computation and Simulation. 2015. V. 85. No. 16. – P. 3191-3205.

[34] Alizadeh Noughabi H. A new estimator of Kullback-Leibler information and its application in goodness of fit tests // Journal of Statistical Computation and Simulation. 2019. V. 89. No. 10. – P. 1914-1934.

[35] Lehmann E. L. Consistency and unbiasedness of certain nonparametric tests // Ann. Math. Statist. – 1951. – Vol. 22, – No. 1. – P. 165-179.

[36] Rosenblatt M. Limit theorems associated with variants of the von Mises statistic // Ann. Math. Statist. 1952. Vol. 23. – P. 617-623.

[37] Lemeshko B.Y. Tests for homogeneity. Guide on the application. M: INFRA–M, 2017. – 207 p. DOI: 10.12737/22368

[38] Lemeshko B.Yu., Lemeshko S.B., Gorbunova A.A. Application and power of criteria for testing the homogeneity of variances. Part I. Parametric criteria // Measurement Techniques. – 2010. – Vol. 53, No. 3. – P. 237-246.

[39] Lemeshko B.Yu., Lemeshko S.B., Gorbunova A.A. Application and power of criteria for testing the homogeneity of variances. Part II. Nonparametric criteria // Measurement Techniques, Vol. 53, No. 5, 2010. – P. 476-486.

[40] Lemeshko B.Y., Sataeva T.S. Application and Power of Parametric Criteria for Testing the Homogeneity of Variances. Part III // Measurement Techniques, 2017. Vol. 60. No. 1. – P. 7-14.

[41] Lemeshko B.Y., Sataeva T.S. Application and Power of Parametric Criteria for Testing the Homogeneity of Variances. Part IV // Measurement Techniques, 2017. Vol. 60. No. 5. – P. 425-431.

[42] Cochran W. G. The distribution of the largest of a set of estimated variances as a fraction of their total // Annals of Eugenics. 1941. Vol. 11. – P. 47-52.

[43] Klotz J. Nonparametric tests for scale // Ann. Math. Stat. 1962. Vol. 33. – P. 498-512.

On Application of k-samples homogeneity tests

Boris Yu. Lemeshko and Irina V. Veretelnikova
Novosibirsk State Technical University, Novosibirsk, Russian Federation
e-mail: Lemeshko@ami.nstu.ru, ira-veterok@mail.ru

Abstract

New k-samples homogeneity tests based on the Smirnov, Lehmann-Rosenblatt and Anderson-Darling two-sample tests have been proposed. The maximum value of the statistics of the 2-sample test obtained during the analysis of combinations of pairs of samples is considered as a statistic of k-sample test. The constructed models for limit distributions of statistics of the proposed tests for $k = 3, \cdots, 11$ are given. Comparative analysis of the power of the set of k-samples tests, including the Zhang test, has been carried out. Power estimates of the studied tests are presented in relation to some competing hypotheses, which allows to order k-sample tests by preference with respect to different alternatives.

Keywords: k-samples tests, homogeneity tests, test statistic, distribution of statistics, power of test.

Introduction

The necessity of solving the task of checking the hypotheses of two (or more) samples of random values belonging to the same universe estimates (the homogeneity test) may arise in different areas. For example, this task may arise naturally when checking the measurement means and trying to be certain that the random measurement errors distribution law has not undergone any serious changes within some time period.

The task of testing the homogeneity of k-samples can be stated as follows. We have x_{ij}, where j is the observation in the set of order statistics of i-sample $j = \overline{1, n_i}, i = \overline{1, k}$. Let us assume that the i-sample correlates with the continuous distribution function of $F_i(x)$. It is required to test the hypothesis of $H_0 : F_1(x) = F_2(x) = \cdots = F_k(x)$ type without defining the common distribution law.

The general approach to the construction of k-sample homogeneity tests which are the counterparts of the two-sample Kolmogorov-Smirnov and Cramer-von Mises (Lehmann-Rosenblatt) tests, was considered in [1]. Under this approach, the statistics of the criterion is a measure of deviation of empirical distributions corresponding to specific samples from the empirical distribution based on the totality of the analyzed samples. The k-selective variant of the Kolmogorov–Smirnov test based on this principle is mentioned in [2, 3]. The k-selective version of the Anderson-Darling test is proposed in [4]. The homogeneity tests constructed by Zhang in [5, 6, 7] are the development of the homogeneity tests by Smirnov [8], Lehmann-Rosenblatt [9, 10] and Anderson–Darling [11] and allow us to analyze samples.

The application of k-samples tests in practice is constrained by the fact that, at best, only critical values of statistics for the relevant ones are known, as in the case of the Anderson-Darling test [4] or Kolmogorov-Smirnov tests [2, 12], and the possibility

of using Zhang's criteria rests on the need to look for the distribution of test statistics (or estimation of the achieved significance level p_{value}) using statistical modeling in order to form a conclusion about the results of the hypothesis test.

The only exception is the homogeneity test χ^2 for which the asymptotic distributions of statistics are known with the truth of H_0.

In the present work we illustrate the dependence of the distributions of statistics of the k-sample tests on the sample sizes and the number of k compared samples. For the k-sample Anderson–Darling test [4] we give models of limit distributions of statistics constructed by us [13, 14, 15]. Suggested variants of k-sample tests based on the use of 2-sample Smirnov test [8], Lehmann-Rosenblatt test [9, 10] and Anderson-Darling test [11], and present the constructed model for the limit distributions of the statistics of the proposed test for various k. The constructed models make it possible to carry out correct and informative conclusions with the calculation of p_{value} with the usage of the corresponding criteria. In addition, we present estimates of the power of the test considered with respect to some competing hypotheses, which allows us to organize the k-sample tests by preference with respect to various alternatives.

The studies were based on the intensive use of the Monte Carlo method in the simulation of distributions of tests statistics.

1 k-samples homogeneity tests

1.1 Anderson-Darling test

The Anderson-Darling k-sample test is proposed in [4]. Let us denote the empirical distribution function corresponding to the i^{th} sample $F_{in_i}(x)$, and the empirical distribution function corresponding to the combined sample volume $n = \sum_{i=1}^{k} n_i$ as $H_n(x)$. Statistics of the Anderson-Darling sample test (AD) is defined by the expression

$$A^2_{kn} = \sum_{i=1}^{k} n_i \int_{B_n} \frac{[F_{in_i}(x) - H_n(x)]^2}{(1 - H_n(x))H_n(x)} dH_n(x),$$

where $B_n = x \in R : H_n(x) < 1$. Under the assumption of continuity of $F_i(x)$ on the ordered combined sample $X_1 \leq X_2 \cdots = X_n$ in [4] this simple expression for the calculation of statistics is obtained:

$$A^2_{kn} = \frac{1}{n} \sum_{i=1}^{k} \frac{1}{n_i} \sum_{j=1}^{n-1} \frac{(nM_{ij} - jn_i)^2}{j(n-j)},$$

where M_{ij} is number of elements in i^{th} sample which are not larger than X_j. The hypothesis H_0 being tested is rejected for large values of statistics.

The statistics acquires the following final form in [4]:

$$T_{kn} = \frac{A^2_{kn} - (k-1)}{\sqrt{D[A^2_{kn}]}}. \tag{1}$$

where the dispersion is determined by the following expression [4]

$$D[A_{kn}^2] = \frac{an^3+bn^2+cn+d}{(n-1)(n-2)(n-3)}$$

with

$$\begin{gathered} a = (4g-6)(k-1) + (10-6g)H, \\ b = (2g-4)k^2 + 8hk + (2g-14h-4)H - 8h + 4g - 6, \\ c = (6h+2g-2)k^2 + (4h-4g+6)k + (2h-6)H + 4h, \\ d = (2h+6)k^2 - 4hk, \end{gathered}$$

where

$$H = \sum_{i=1}^{k} \frac{1}{n_i}, h = \sum_{i=1}^{n-1} \frac{1}{i}, g = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \frac{1}{(n-i)j}.$$

Asymptotic (limiting) distributions of statistics (1) depend on the k-number of samples compared and do not depend on n_i. With the growth of k the distribution of statistics (1) slowly converges to the standard normal law.

In [4] for statistics (1) the table of critical values has been constructed for a number of k. Based on the results of statistical modeling, we built models of limiting distributions of statistics (1) for [13, 14, 15]. The laws of the family of beta-distributions of the III type with density turned out to be good models when having the density of

$$f(x) = \frac{\theta_2^{\theta_0}}{\theta_3 B(\theta_0, \theta_1)} \left[\frac{x-\theta_4}{\theta_3}\right]^{\theta_0-1} \left[1 - \frac{x-\theta_4}{\theta_3}\right]^{\theta_1-1} / \left[1 + (\theta_2 - 1)\frac{x-\theta_4}{\theta_3}\right]^{\theta_0+\theta_1}, \quad (2)$$

as shown in Table 1 as $B_{III}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ having exact values for this law's parameters. These models are based on simulated samples of statistics with the number of simulation experiments $N = 10^6$ and $n_i = 10^3$.

1.2 Zhang test

The Zhang tests [5, 6, 7] allow comparing $k \geq 2$ samples.

Let $x_{i1}, x_{i2}, \cdots, x_{in_i}$ be ordered samples of continuous random variables with distribution functions $F_i(x)$, $(i = \overline{1,k})$ and, as previously, $X_1 < X_2 < \cdots < X_n$, where $n = \sum_{i=1}^{k} n_i$, is the unified ordered sample. Let us define the R_{ij} rank of the j^{th} ordered observation x_{ij} of the i^{th} sample in the unified sample. Let $X_0 = -\infty, X_{n+1} = +\infty$, and the ranks $R_{i,0} = 1, R_{i,n_i+1} = n+1$.

In the tests a modification of the empirical distribution function $\hat{F}(t)$ is used, having the values of $\hat{F}(X_m) = (m-0.5)/n$ at break points $X_m, m = \overline{1,n}$ [5].

The Z_K statistic of the Zhang homogeneity test is of the following form [5]:

$$Z_K = \max_{1 \leq m \leq n} \left\{ \sum_{i=1}^{k} n_i \left[F_{i,m} \ln \frac{F_{i,m}}{F_m} + (1 - F_{i,m}) \ln \frac{1-F_{i,m}}{1-F_m} \right] \right\}, \quad (3)$$

Table 1: Models of the limiting distributions of statistics (1)

k	Model
2	B_{III}(3.1575, 2.8730, 18.1238, 15.0000, -1.1600)
3	B_{III}(3.5907, 4.5984, 7.8040, 14.1310, -1.5000)
4	B_{III}(4.2657, 5.7035, 5.3533, 12.8243, -1.7500)
5	B_{III}(6.2992, 6.5558, 5.6833, 13.010, -2.0640)
6	B_{III}(6.7446, 7.1047, 5.0450, 12.8562, -2.2000)
7	B_{III}(6.7615, 7.4823, 4.0083, 11.800, -2.3150)
8	B_{III}(5.8057, 7.8755, 2.9244, 10.900, -2.3100)
9	B_{III}(9.0736, 7.4112, 4.1072, 10.800, -2.6310)
10	B_{III}(10.2571, 7.9758, 4.1383, 11.186, -2.7988)
11	B_{III}(10.6848, 7.5950, 4.2041, 10.734, -2.8400)
∞	N(0.0,1.0)

where $F_m = \hat{F}(X_m)$, so that $F_m = (m-0.5)/n$, and the calculation $F_{i,m} = \hat{F}_i(X_m)$ is done as follows. At the initial moment $j_i = 0, i = \overline{1,k}$. If $R_{i,j_i+1} = m$, then $j_i := j_i+1$ and $F_{i,m} = (j_i - 0.5)/n_i$, otherwise, with $R_{i,j_i} < m < R_{i,j_i+1}$, $F_{i,m} = j_i/n_i$.

This is a *right-hand* test: the hypothesis H_0 being tested is rejected at *high* statistical values (3).

Statistic Z_A of the homogeneity test of k samples is defined by the following expression [5]:

$$Z_A = -\sum_{m=1}^{n}\sum_{i=1}^{k} n_i \frac{F_{i,m}\ln F_{i,m} + (1-F_{i,m})\ln(1-F_{i,m})}{(m-0.5)(n-m+0.5)}, \tag{4}$$

where F_m and $F_{i,m}$ are calculated as shown above.

This is a *left-side* test: the hypothesis H_0 being tested is rejected for *small* values of statistics (4).

Distributions of the statistic (4) depend on the sample volume and the number of samples compared as well.

Statistic Z_C of the homogeneity test of k samples is defined by the following expression [5]:

$$Z_C = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} \ln\left(\frac{n_i}{j-0.5} - 1\right)\ln\left(\frac{n}{R_{i,j}-0.5} - 1\right). \tag{5}$$

This is also a *left-hand* test: the tested hypothesis H_0 is rejected at *small* values of the statistic (5). The distributions $G(Z_C \mid H_0)$ of the statistic depend on the sample volume and the number of samples under analysis in the similar way.

The dependence of the distributions of statistics (3) - (5) of the volume of the samples complicates the use of the Zhang test since there are problems with the calculation of the evaluation of p_{value}.

At the same time, the lack of information on the laws of distribution of statistics and tables of critical values in modern conditions is not a serious disadvantage of the tests as it is easy to calculate the achieved levels of significance of p_{value} with the software that supports the application of the tests, merely using statistic simulating methods.

1.3 k-samples Tests Based on 2-sample Ones

In order to analyze the k-samples it is possible to apply a two-sample test with the S statistic to each pair (totaling $(k-1)k/2$ pairs), and the decision on accepting or rejecting the H_0 hypothesis will be made on the strength of all results. The following statistic can be taken as a statistic of this k-sample tests (when having a right-hand two-sample criterion):

$$S_{max} = \max_{\substack{1\le i\le k\\ i<j\le k}} \{S_{ij}\}, \tag{6}$$

where S_{ij} are the values of the statistics of the used two-sample criterion as calculated in the course of analysis of the i^{th} and the j^{th} samples.

The hypothesis H_0 to be tested will be rejected at **large** values of statistics S_{max}. The advantage of this kind of test is that as a result a pair of samples will be determined, the difference between them being the most significant from the standpoint of the two-sample test used.

Statistics of the two-sample Smirnov, Lehmann-Rosenblatt and Anderson-Darling tests can be used as S_{ij}. In this case the distributions of the relevant statistics S_{max} converge to some limiting ones, models of which can be found on the results of statistical modeling.

1.3.1 Smirnov Maximum Test

The D_{n_2,n_1} statistic used in the Smirnov test is calculated according to the following formulae [8]:

$$D^+_{n_2,n_1} = \max_{1\le r\le n_2} [\tfrac{r}{n_2} - F_{1,n_1}(x_{2r})] = \max_{1\le s\le n_1} [F_{2,n_2}(x_{2s}) - \tfrac{s-1}{n_1}],$$

$$D^-_{n_2,n_1} = \max_{1\le r\le n_2} [F_{1,n_1}(x_{2r}) - \tfrac{r-1}{n_2}] = \max_{1\le s\le n_1} [\tfrac{s}{n_1} - F_{2,n_2}(x_{1s})],$$

$$D_{n_2,n_1} = \max(D^+_{n_2,n_1}, D^-_{n_2,n_1}).$$

With the H_0 hypothesis being true and with unlimited increase of the number of samples the statistic

$$S_C = \sqrt{\frac{n_1 n_2}{n_1+n_2}} D_{n_2,n_1} \tag{7}$$

will in the limit fall with the Kolmogorov arrangement of $K(S)$ [8].

In case of using the k-samples variant of the Smirnov test as S_{ij} in (6) it seems more preferable to use a modification of the Smirnov statistic

$$S_{mod} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}\left(D_{n_2,n_1} + \frac{n_1 + n_2}{4.6 n_1 n_2}\right), \tag{8}$$

its distribution being always closer to the limiting distribution of Kolmogorov $K(S)$ [16]. Statistic S_{max} will be defined as $S_{max}^{S_m}$ in this case.

With equal volumes of samples under comparison the statistic distributions $S_{max}^{S_m}$ will be of substantial discreteness (similar to the two-sample case, see Fig. 1) and be different from the asymptotic (limiting) distributions (see Fig. 2). If possible, it is preferable to use co-primes as n_i, then the distributions $G(S \mid H_0)$ of the $S_{max}^{S_m}$ statistic will not be actually different from the asymptotic ones.

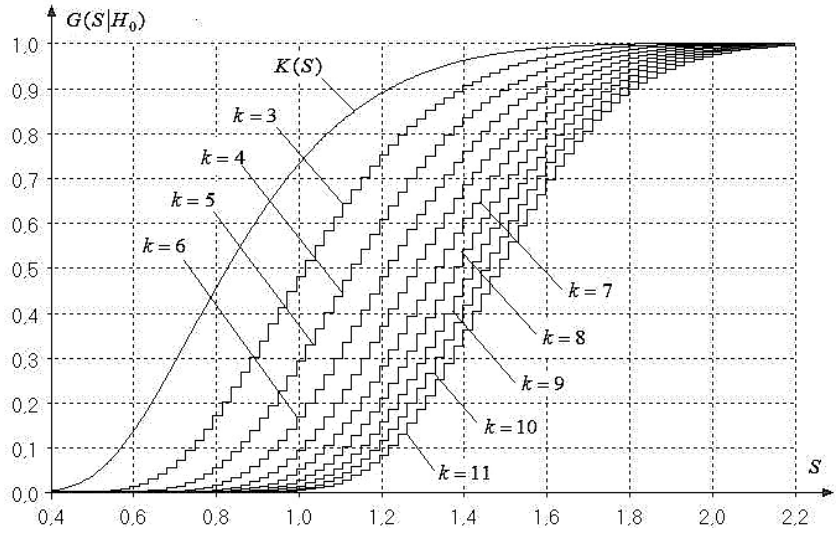


Figure 1: Statistic distributions with $n_i = 1000, i = \overline{1,k}$

Models of asymptotic $S_{max}^{S_m}$ statistic distributions with $k = 3 \div 11$ in the form of beta distributions of the III type (2) $B_{III}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ having exact values of parameters and constructed in this paper based on the results of statistic modeling are shown in Table 2.

1.3.2 Lehman-Rosenblatt Maximum Test

Statistic of the two-sample Lehmann-Rosenblatt test as introduced in [9] is used in the following form [8]:

$$T = \frac{1}{n_1 n_2 (n_1 + n_2)}\left(n_2 \sum_{i=1}^{n_2} (r_i - i)^2 + n_1 \sum_{j=1}^{n_1} (s_j - j)^2\right) - \frac{4 n_1 n_2 - 1}{6(n_1 + n_2)}, \tag{9}$$

where r_i is the numerical order (rank) of x_{2i}; s_j is the numerical order (rank) of x_{1i} in the unified ordered series. In [10] it was shown that the statistic (9) at the limit is distributed as $a1(t)$ [8].

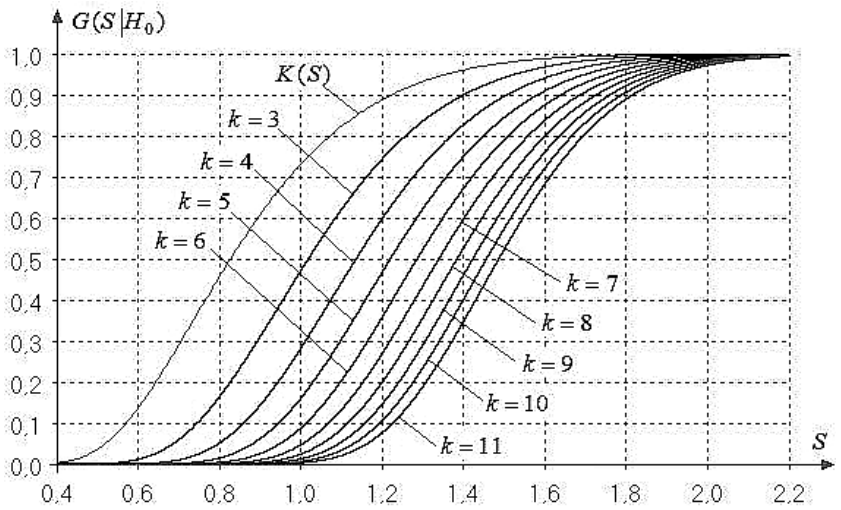


Figure 2: Asymptotic statistic distributions $S_{max}^{S_m}$

Table 2: Models of the limiting distributions of statistics $S_{max}^{S_m}$

k	Model
2	$K(S)$
3	B_{III}(6.3274, 6.6162, 2.8238, 2.4073, 0.4100)
4	B_{III}(7.2729, 7.2061, 2.6170, 2.3775, 0.4740)
5	B_{III}(7.1318, 7.3365, 2.4813, 2.3353, 0.5630)
6	B_{III}(7.0755, 8.0449, 2.3163, 2.3818, 0.6320)
7	B_{III}(7.7347, 8.6845, 2.3492, 2.4479, 0.6675)
8	B_{III}(7.8162, 8.9073, 2.2688, 2.4161, 0.7120)
9	B_{III}(7.8436, 8.8805, 2.1696, 2.3309, 0.7500)
10	B_{III}(7.8756, 8.9051, 2.1977, 2.3280, 0.7900)
11	B_{III}(7.9122, 9.0411, 2.1173, 2.2860, 0.8200)

In the case of using the k-samples variant of the Lehman-Rosenblatt test as S_{ij} in the statistic S_{max}^{LR} of form (6) statistic (9) is used. Dependence of distributions of statistic S_{max}^{LR} on the number of samples with H_0 being true is illustrated in Fig. 3.

The constructed models of asymptotic (limiting) distributions of statistic S_{max}^{LR} with the number of compared samples $k = 3 \div 11$ are shown in Table 3. In this case the Sb-Johnson distributions proved to be the best with the density of

$$f(x) = \frac{\theta_1\theta_2}{\sqrt{2\pi}(x-\theta_3)(\theta_2+\theta_3-x)} \exp\left\{-\frac{1}{2}\left[\theta_0 - \theta_1 \ln \frac{x-\theta_3}{\theta_2+\theta_3-x}\right]^2\right\}$$

with exact values of this law's parameters, the law being shown in Table 3 as $Sb(\theta_0, \theta_1, \theta_2, \theta_3)$. These represented models allow finding the estimates of p_{value} by the values of statistic S_{max}^{LR} with corresponding k number of samples under comparison.

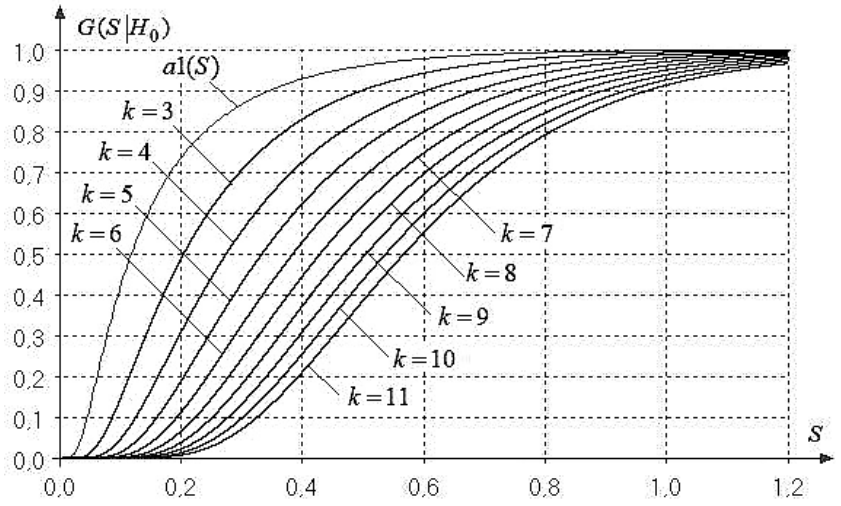


Figure 3: Distributions of statistic S_{max}^{LR}

Table 3: Models of the limiting distributions of statistics S_{max}^{LR}

k	Model
2	$a1(t)$
3	$Sb(3.2854, 1.2036, 3.0000, 0.0215)$
4	$Sb(2.5801, 1.2167, 2.2367, 0.0356)$
5	$Sb(3.1719, 1.4134, 3.1500, 0.0320)$
6	$Sb(2.9979, 1.4768, 2.9850, 0.0380)$
7	$Sb(3.2030, 1.5526, 3.4050, 0.0450)$
8	$Sb(3.2671, 1.6302, 3.5522, 0.0470)$
9	$Sb(3.4548, 1.7127, 3.8800, 0.0490)$
10	$Sb(3.4887, 1.7729, 3.9680, 0.0510)$
11	$Sb(3.4627, 1.8168, 3.9680, 0.0544)$

1.3.3 Anderson-Darling Maximum Test

The Anderson-Darling two-sample test was dealt with in [11]. This test's statistic is defined by the following expression:

$$A^2 = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1+n_2-1} \frac{(M_i(n_1+n_2) - n_1 i)^2}{i(n_1+n_2-i)}, \tag{10}$$

where M_i is the number of elements of the first sample, smaller or equal to the i^{th} element of the variation set of the unified sample. Distribution $a2(t)$ will be the limiting distribution (10) with the tested hypothesis H_0 being true [8].

In the case of using the k-samples variant of the Anderson-Darling test as S_{ij} in the S_{max}^{AD} statistic (6) statistic (10) will be used. Dependence of distributions of

statistic S^{AD}_{max} on the number of samples with H_0 being true is shown in Fig. 4.

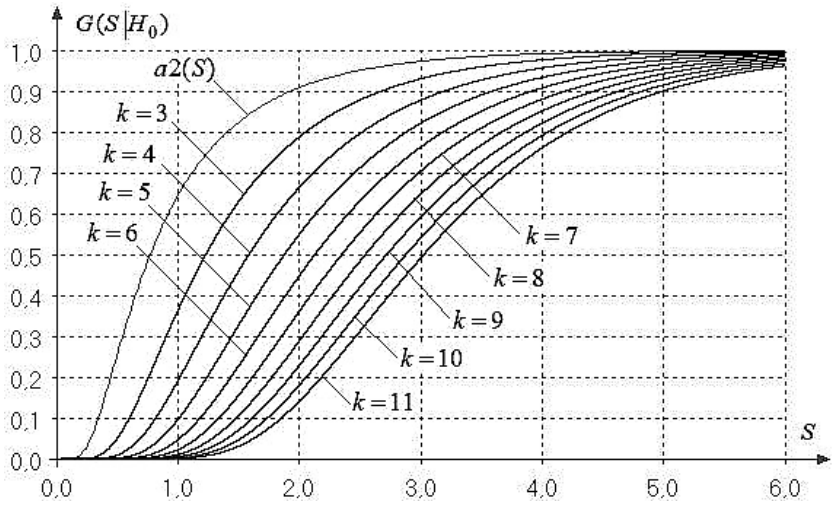


Figure 4: Distributions of statistic S^{AD}_{max}

Models of asymptotic (limiting) distributions of statistic S^{AD}_{max} for the k number of samples under comparison $k = 3 \div 11$ have been constructed for distributions $G(S^{AD}_{max} \mid H_0)$ and shown in Table 4. In this case the beta distributions of the III type proved to be the best (2) as shown as $B_{III}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ with exact values of parameters shown in Table 4; these can be used for estimating p_{value} with the k number of compared samples.

Table 4: Models of the limiting distributions of statistics S^{AD}_{max}

k	Model
2	$a2(t)$
3	B_{III}(4.4325, 2.7425, 12.1134, 8.500, 0.1850)
4	B_{III}(5.2036, 3.2160, 10.7792, 10.000, 0.2320)
5	B_{III}(5.7527, 3.3017, 9.7365, 10.000, 0.3000)
6	B_{III}(5.5739, 3.4939, 7.7710, 10.000, 0.3750)
7	B_{III}(6.4892, 3.6656, 8.0529, 10.500, 0.3920)
8	B_{III}(6.3877, 3.8143, 7.3602, 10.800, 0.4800)
9	B_{III}(6.7910, 3.9858, 7.1280, 11.100, 0.5150)
10	B_{III}(6.7533, 4.2779, 6.6457, 11.700, 0.5800)
11	B_{III}(7.1745, 4.3469, 6.6161, 11.800, 0.6100)

1.4 Homogeneity Test χ^2

The homogeneity test χ^2 can successfully be used to analyze $k \geq 2$ samples. In this case the common area of the samples is split into r intervals (groups). Let η_{ij} be the

number of elements of the i^{th} sample of the j^{th} interval, then $n_i = \sum_{j=1}^{r} \eta_{ij}$.

The χ^2 homogeneity test statistic will be of the following form:

$$\chi^2 = n \sum_{i=1}^{k} \sum_{j=1}^{r} \frac{(\eta_{ij} - \nu_j n_i/n)^2}{\nu_j n_i} = n\left(\sum_{i=1}^{k} \sum_{j=1}^{r} \frac{\eta_{ij}^2}{\nu_j n_i} - 1 \right), \tag{11}$$

where $\nu_j = \sum_{l=1}^{k} \eta_{lj}$ is the total number of elements of all samples falling into the j^{th} interval. The χ^2-distribution with the number of degrees of freedom $(k-1)(r-1)$ shall be the asymptotic distribution of statistic [17].

2 Comparative analysis of powers

One of the main characteristics of the statistical test is its power relative to a given competing hypothesis H_1. The power is the remainder of $1-\beta$, where β is the possibility of type II error (accept hypothesis H_0 with H_1 being true) at specified probability α of type I error (reject H_0 when true).

The power of k-samples tests was investigated for various k and situations when the tested hypothesis H_0 was whether all samples belonged to the standard normal law, the competing hypothesis H_1 being if all samples but the last one belonged to the standard normal law and the last sample belonged to the normal law with the shift parameter $\theta_0 = 0.1$ and the scale parameter $\theta_1 = 1$; hypothesis H_2 being that the last sample belonged to the normal law with the shift parameter $\theta_0 = 0$ and the scale parameter $\theta_1 = 1.1$, the competing hypothesis H_3 being the last sample belonged to the logistic law with the density of

$$f(x) = \tfrac{1}{\theta_1\sqrt{3}} \exp\{-\tfrac{\pi(x-\theta_0)}{\theta_1\sqrt{3}}\}/[1+\exp\{-\tfrac{\pi(x-\theta_0)}{\theta_1\sqrt{3}}\}]^2$$

and parameters $\theta_0 = 0$ and $\theta_1 = 1$.

The power was evaluated on the results of modeling statistic distributions with the tested $G(S \mid H_0)$ being true, and competing hypotheses $G(S \mid H_1)$, $G(S \mid H_2)$ and $G(S \mid H_3)$ having equal volumes of n_i compared samples. As an example, Tables 5 and 6 show evaluation of the power of tests with $\alpha = 0.1$ for $k = 3$ and $k = 4$ correspondingly. In the case of the homogeneity test χ^2 the unified sample was split into $r = 10$ equifrequent intervals.

Thus-conducted power analysis of k-samples tests allows making some conclusions.

The tests can be organized power-wise with respect to changes in the shift parameter in the following way:

$$S_{max}^{AD} \succ AD \succ S_{max}^{LR} \succ S_{max}^{Sm} \succ Z_C \succ Z_A \succ Z_K \succ \chi^2.$$

With respect to changes in the scale parameter:

$$Z_C \succ Z_A \succ Z_K \succ AD \succ \chi^2 \succ S_{max}^{AD} \succ S_{max}^{Sm} \succ S_{max}^{LR}.$$

Table 5: Assessment of the power of test against alternatives H_1, H_2 and H_3, $k = 3$, $n_i = n$

Test	$n_i = 20$	$n_i = 50$	$n_i = 100$	$n_i = 300$	$n_i = 500$	$n_i = 10^3$
Against alternative hypothesis H_1						
S^{AD}_{max}	0.113	0.134	0.171	0.314	0.450	0.712
AD	0.113	0.134	0.171	0.313	0.449	0.711
S^{LR}_{max}	0.114	0.134	0.168	0.306	0.437	0.694
S^{Sm}_{max}	0.110	0.128	0.155	0.272	0.383	0.622
Z_C	0.113	0.131	0.160	0.273	0.380	0.612
Z_A	0.112	0.130	0.158	0.268	0.371	0.599
Z_K	0.110	0.125	0.144	0.231	0.321	0.525
χ^2	0.100	0.108	0.120	0.173	0.226	0.385
Against alternative hypothesis H_2						
Z_C	0.107	0.125	0.160	0.319	0.475	0.771
Z_A	0.107	0.126	0.162	0.319	0.470	0.767
Z_K	0.107	0.123	0.147	0.263	0.376	0.621
AD	0.104	0.111	0.124	0.191	0.273	0.509
χ^2	0.105	0.114	0.129	0.202	0.277	0.495
S^{AD}_{max}	0.102	0.107	0.114	0.165	0.231	0.446
S^{Sm}_{max}	0.103	0.104	0.114	0.136	0.164	0.253
S^{LR}_{max}	0.103	0.104	0.108	0.127	0.152	0.241
Against alternative hypothesis H_3						
Z_A	0.103	0.108	0.116	0.181	0.279	0.580
Z_C	0.103	0.108	0.116	0.176	0.270	0.568
Z_K	0.104	0.110	0.117	0.170	0.233	0.423
χ^2	0.100	0.113	0.121	0.173	0.226	0.382
AD	0.103	0.107	0.114	0.148	0.189	0.315
S^{Sm}_{max}	0.102	0.105	0.111	0.148	0.183	0.288
S^{AD}_{max}	0.102	0.104	0.110	0.134	0.166	0.272
S^{LR}_{max}	0.103	0.104	0.107	0.124	0.145	0.218

At that, the Zhang tests of Z_A and Z_C statistics are almost equivalent power-wise, and the Anderson-Darling test is noticeably inferior to the Zhang tests.

The tests can be organized power-wise with respect to situations when all but one sample belongs to the normal law and the last one belongs to the logistic law, in the following way:

$$Z_A \succ Z_C \succ Z_K \succ \chi^2 \succ AD \succ S^{Sm}_{max} \succ S^{AD}_{max} \succ S^{LR}_{max}.$$

It can be noted that with the increase in the number of compared samples of the same volumes the power of the criterion relative to similar competing hypotheses

decreases as a rule, which is absolutely natural. It is more difficult to single out a situation and to give preference to a competing hypothesis, when only one of the analyzed samples belongs to some other law. We can't but mention that the Zhang tests with statistics of Z_K, Z_A, Z_C possess quite substantial advantage in power with respect to some alternatives.

Table 6: Assessment of the power of test against alternatives H_1, H_2 and H_3, $k = 4$, $n_i = n$

Test	$n_i = 20$	$n_i = 50$	$n_i = 100$	$n_i = 300$	$n_i = 500$	$n_i = 10^3$
Against alternative hypothesis H_1						
S_{max}^{AD}	0.112	0.131	0.165	0.302	0.438	0.706
AD	0.112	0.131	0.164	0.301	0.433	0.701
S_{max}^{LR}	0.113	0.130	0.162	0.293	0.425	0.686
S_{max}^{Sm}	0.111	0.125	0.151	0.261	0.366	0.605
Z_C	0.111	0.126	0.155	0.260	0.368	0.595
Z_A	0.111	0.127	0.153	0.255	0.360	0.579
Z_K	0.109	0.121	0.141	0.219	0.300	0.502
χ^2	0.102	0.109	0.118	0.167	0.221	0.358
Against alternative hypothesis H_2						
Z_C	0.106	0.122	0.158	0.306	0.468	0.761
Z_A	0.107	0.124	0.158	0.305	0.463	0.745
Z_K	0.106	0.120	0.145	0.249	0.367	0.606
AD	0.104	0.110	0.123	0.180	0.254	0.474
χ^2	0.107	0.113	0.127	0.189	0.271	0.458
S_{max}^{AD}	0.101	0.104	0.111	0.145	0.195	0.381
S_{max}^{Sm}	0.102	0.105	0.108	0.128	0.153	0.221
S_{max}^{LR}	0.102	0.103	0.105	0.118	0.135	0.197
Against alternative hypothesis H_3						
Z_A	0.103	0.107	0.116	0.179	0.274	0.566
Z_C	0.103	0.107	0.115	0.173	0.257	0.555
Z_K	0.103	0.107	0.114	0.161	0.222	0.410
χ^2	0.102	0.110	0.116	0.164	0.218	0.357
AD	0.102	0.106	0.113	0.143	0.179	0.291
S_{max}^{Sm}	0.103	0.104	0.112	0.138	0.166	0.257
S_{max}^{AD}	0.101	0.103	0.107	0.124	0.147	0.229
S_{max}^{LR}	0.102	0.102	0.105	0.116	0.130	0.183

Conclusions

The constructed models of statistic limiting distributions for k-samples homogeneity tests (the Anderson-Darling ones and those proposed in this paper) allows obtaining correct and informational conclusions on and calculating the tests significance p_{value}. Software can is available for this purpose [18].

Funding

The studies were carried out with the support of the Ministry of Education and Science of the Russian Federation in the framework of the state work 'Ensuring the conduct of scientific research'(No. 1.4574.2017 / 6.7) and the design part of the state task (No. 1.1009.2017 / 4.6).

References

[1] Kiefer J. (1959). K-samples Analogues of the Kolmogorov–Smirnov and Cramer-v. Mises Tests. *Annals of Mathematical Statistics*. Vol. **30. No. 2.**, – pp. 420-447. URL: http://www.jstor.org/ stable/2237091

[2] Conover W.J.(1965). Several k-samples Kolmogorov–Smirnov tests. *The Annals of Mathematical Statistics*. Vol. **36, No. 3**, pp. 1019-1026. URL: http://www.jstor.org/ stable/2238210

[3] Conover W.J. (1999). *Practical Nonparametric Statistics – 3d ed.*. Wiley.

[4] Scholz F.W., Stephens M.A. (1987). K-samples Anderson–Darling Tests. *Journal of the American Statistical Association*. Vol. **82. No. 399**, pp. 918-924. DOI: 10.1080/ 01621459.1987.10478517

[5] Zhang J. (2001). *Powerful goodness-of-fit and multi-sample tests*. PhD Thesis. York University, Toronto. URL: http://www.collectionscanada.gc.ca/ obj/s4/f2/dsk3/ ftp05/ NQ66371.pdf (accessed 28.01.2013).

[6] Zhang J. (2006). Powerful Two-Sample Tests Based on the Likelihood Ratio. *Technometrics*. Vol. **48. No. 1**, pp. 95-103. DOI: 10.1198/ 004017005000000328

[7] Zhang J., Wu Y. (2007). k-samples tests based on the likelihood ratio. *Computational Statistics & Data Analysis*. Vol. **51. No. 9**, pp. 4682-4691. DOI: 10.1016/ j.csda.2006.08.029

[8] Bolshev L.N., Smirnov N.V. (1983).*Tables of mathematical statistic*. M.: Nauka.

[9] Lehmann E.L. (1951). Consistency and unbiasedness of certain nonparametric tests. *Ann. Math. Statist.* Vol. **22, № 1**,pp. 165–179. URL: http://www.jstor.org/stable/2236420

[10] Rosenblatt M. (1952). Limit theorems associated with variants of the von Mises statistic. *Ann. Math. Statist.* Vol. **23**, pp. 617–623. URL: https://projecteuclid.org/euclid.aoms/1177729341

[11] Pettitt A.N. (1976). A two-sample Anderson-Darling rank statistic.*Biometrika*. Vol. **63. No.1**, pp. 161-168. DOI: 10.1093/biomet/63.1.161

[12] Wolf E.H., Naus J.I. (1973). Tables of critical values for a k-sample Kolmogorov–Smirnov test statistic.*J. Amer. Statist. Assoc.* Vol. **68**, pp. 994–997. DOI: 10.1080/01621459.1973.10481462

[13] Lemeshko B.Y. (2017).*Tests for homogeneity. Guide on the application*. M: INFRA–M. DOI: 10.12737/22368

[14] Lemeshko B.Y., Lemeshko S.B., Veretelnikova I.V. (2017). On application of distribution laws homogeneity tests.*Tomsk State University Journal of Control and Computer Science*. No. **41**, pp. 24-31. DOI: 10.17223/19988605/41/3

[15] Lemeshko B.Y., Veretelnikova I.V., Lemeshko S.B., Novikova A.Y. (2017). Application of Homogeneity Tests: Problems and Solution.*In: Rykov V., Singpurwalla N., Zubkov A. (eds) Analytical and Computational Methods in Probability Theory. ACMPT 2017. Lecture Notes in Computer Science*. Vol. **10684**.

[16] Lemeshko B.Y., Lemeshko S.B. (2005). Statistical distribution convergence and homogeneity test power for Smirnov and Lehmann–Rosenblatt tests.*Measurement Techniques*. Vol. **48,No.12**, pp. 1159-1166. DOI: 10.1007/s11018-006-0038-3

[17] Cramer H. (1975).*Mathematical methods of statistics*. M.: Mir.

[18] ISW-Software for statistic analysis of one-dimensional observations. https://ami.nstu.ru/ headrd/ISW.htm. (accessed date 06.06.2019)

On conditional optimization of "kernel" estimators of densities

Anton V. Voytishek[1,2] and Tatyana E. Bulgakova[2]
[1] *Institute of Computational Mathematics and Mathematical Geophysics SD RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Russia*
e-mail: vav@osmf.sscc.ru, tatyana.bulgakova@gmail.com

Abstract

In this paper, it is shown that the "kernel" algorithm for approximation of probability densities, which includes the considerations of numerical mesh approximation of functions, is nearly equal to the randomized projection-mesh functional numerical algorithm of the multi-dimensional analogue of the polygon of frequencies method for approximation of solution of integral Fredholm equation of the second kind. It means that the considerations of the conditional optimization theory of the multi-dimensional analogue of the polygon of frequencies method can be used for the "kernel" algorithm for approximation of probability densities.

Keywords: the randomized projection-mesh functional numerical algorithms, the multi-dimensional analogue of the polygon of frequencies method, the "kernel" estimators for approximation of probability densities, numerical mesh approximation of functions

1. The multi-dimensional analogue of the polygon of frequencies method

In recent years, the theory of randomized functional algorithms is developed (especially in Novosibirsk scientific school of Monte Carlo methods); see, in particular, [1–4]. The most informative examples of these algorithms are related to approximation of the unknown solution $\varphi(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$ of the integral Fredholm equation of the second kind

$$\varphi(\mathbf{x}) = \int k(\mathbf{x}', \mathbf{x})\varphi(\mathbf{x}')\, d\mathbf{x}' + f(\mathbf{x}) \quad \text{or} \quad \varphi = K\varphi + f, \tag{1.1}$$

in a bounded domain $X \subset \mathbb{R}^d$; here $k(\mathbf{x}', \mathbf{x})$ (the kernel of the integral operator K) and $f(\mathbf{x})$ (the free term of the equation) are given functions.

For approximation of the function $\varphi(\mathbf{x})$ we use the representations of classical theory of numerical function approximation (see, for example, [5]), which have the common form

$$\varphi(\mathbf{x}) \approx L^{(M)}\varphi(\mathbf{x}) = \sum_{i=1}^{M} w^{(i)}\chi^{(i)}(\mathbf{x}) \tag{1.2}$$

for some specially selected set of basic functions

$$\Xi^{(M)} = \left\{\chi^{(1)}(\mathbf{x}), ..., \chi^{(M)}(\mathbf{x})\right\}, \tag{1.3}$$

(the form of these functions defines the type of the approximation (1.2)) and coefficients

$$\mathbf{W}^{(M)} = \left\{w^{(1)}, ..., w^{(M)}\right\}, \tag{1.4}$$

which are defined as functionals of the unknown approximated function $\varphi(\mathbf{x})$.

For the randomized functional algorithms, the coefficients (1.4) are calculated approximately using the Monte Carlo method with the test numbers n_i: $w^{(i)} \approx \tilde{w}^{(i)}(n_i)$ (in this paper we investigate the case $n_1 = ... = n_M \equiv n$), and the approximation

$$\varphi(\mathbf{x}) \approx L^{(M)}\tilde{\varphi}(\mathbf{x}) = \sum_{i=1}^{M} \tilde{w}^{(i)}(n)\chi^{(i)}(\mathbf{x}) \tag{1.5}$$

is considered.

In the recent papers [6, 7] we have proposed the new (to compare with the works [1–4]) classification of the randomized functional algorithms for approximation of the solution $\varphi(\mathbf{x})$ of the equation (1.1). We have distinguished *the mesh, the projection and the projection-mesh algorithms* (the type of a method is defined by the choice of the basic functions (1.3) and the coefficients (1.4)). In these papers, we also have presented the considerations why the mesh and the projection randomized functional algorithms can be non-effective (and even unrealizable) for solution of practically important problems related to solutions of integral equations of the form (1.1). In particular, for the theoretically attractive *mesh dependent test method*, the smoothness of the kernel $k(\mathbf{x}', \mathbf{x})$ of the integral operator K is needed. But most part of kernels in applied problems has the integrable singularities (up to delta-functions) and even can not be calculated explicitly. *The mesh adjoint random walk method* is too numerical laborious because of necessity for numerical simulating of individual set of trajectories of the corresponding applied Markov chains for every node $\mathbf{x}_i$ of the introduced mesh

$$X^{(M)} = \{\mathbf{x}_1, ..., \mathbf{x}_M\} \tag{1.6}$$

in the domain X. *The projection methods* have fairly obvious numerical instability.

The projection-mesh randomized functional algorithms have no such flows. For these methods the basic functions (1.3) and the coefficients
$w^{(i)} = w^{(i)}\left(\boldsymbol{\varphi}^{(M)}\right)$; $\boldsymbol{\varphi}^{(M)} = \{\varphi(\mathbf{x}_1), ..., \varphi(\mathbf{x}_M)\}$ from (1.4) are tied with the mesh (1.6) such that they provide a small margin of the deterministic component of error

$$\delta_{det}^{(\mathbb{B}(X))} = \|\varphi - L^{(M)}\varphi\|_{\mathbb{B}(X)}$$

for the used normalized functional space $\mathbb{B}(X)$, together with stability of the approximation (1.5), which is defined by the relative smallness (proximity to unit) of the Lebesgue constant $\tilde{L} = \sup_{\mathbf{x}\in X} \sum_{i=1}^{M} \left|\chi^{(i)}(\mathbf{x})\right|$ from the ratio

$$\delta_{stoch}^{(\mathbb{C}(X))} = \|L^{(M)}\varphi - L^{(M)}\tilde{\varphi}\|_{\mathbb{C}(X)} \le \tilde{L} \max_{i=1,...,M} \left|w^{(i)}\left(\boldsymbol{\varphi}^{(M)}\right) - w^{(i)}\left(\tilde{\boldsymbol{\varphi}}^{(M)}(n)\right)\right|$$

(see, for example, [2]); here $\tilde{\boldsymbol{\varphi}}^{(M)}(n) = \left(\tilde{\varphi}^{(\mathbf{x}_1)}(n), ..., \tilde{\varphi}^{(\mathbf{x}_M)}(n)\right)$ and $\tilde{\varphi}^{(\mathbf{x}_i)}(n)$ is the Monte Carlo approximation of the value $\varphi(\mathbf{x}_i)$; $i = 1, ..., M$.

In this case the approximations of the Monte Carlo method

$$\tilde{\mathbf{W}}^{(M)} = \left\{\tilde{w}^{(1)}(n), ..., \tilde{w}^{(M)}(n)\right\}$$

of the coefficients (1.4) from the ratio (1.5) have the form

$$\tilde{w}^{(i)}(n) = w^{(i)}\left(\tilde{\boldsymbol{\varphi}}^{(M)}(n)\right), \text{ more often } w^{(i)}\left(\tilde{\boldsymbol{\varphi}}^{(M)}(n)\right) = \tilde{\varphi}^{(\mathbf{x}_i)}(n).$$

In turn, to obtain values $\tilde{\boldsymbol{\varphi}}^{(M)}(n)$ for randomized functional *projection-mesh* algorithms the following special technology (which defines the difference from *mesh* functional algorithms) is used. Choose the finite, having the same shape for all $\{\mathbf{x}_1, ..., \mathbf{x}_M\}$ functions (versions of "kernel" function $\kappa^{(\mathbf{x})}(\mathbf{y})$ for various values of the parameter $\mathbf{x}$ – see the Section 2 of this paper)

$$\mathbf{K}^{(M)} = \left\{\kappa^{(\mathbf{x}_1)}(\mathbf{y}), ..., \kappa^{(\mathbf{x}_M)}(\mathbf{y})\right\}, \tag{1.7}$$

related (as the basic functions (1.3)) to the mesh (1.6) such that

$$\int \varphi(\mathbf{y})\kappa^{(\mathbf{x}_i)}(\mathbf{y})\, d\mathbf{y} \approx \varphi(\mathbf{x}_i); \quad i = 1, ..., M. \tag{1.8}$$

Further we recall the classical (see, for example, Chapter 4 of the textbook [3]) considerations that for approximate calculation of linear functionals of the form

$$I_h = \int \varphi(\mathbf{y})h(\mathbf{y})\, d\mathbf{y} \tag{1.9}$$

on the solution $\varphi(\mathbf{x})$ of the equation (1.1) it is expedient to use *the main estimator* (or *unbiased Monte Carlo collision estimate*):

$$I_h = \mathbf{E}\zeta; \quad \zeta = \sum_{m=0}^{N} Q^{(m)} h\left(\boldsymbol{\xi}^{(m)}\right), \tag{1.10}$$

where

$$\boldsymbol{\xi}^{(0)}, \boldsymbol{\xi}^{(1)}, ..., \boldsymbol{\xi}^{(N)} \tag{1.11}$$

is *the applied Markov chain* (or homogeneous Markov chain terminated with unit probability) with the initial density $\pi(\mathbf{x})$ and the transition function $p(\mathbf{x}', \mathbf{x}) = r(\mathbf{x}', \mathbf{x}) \times \left[1 - p^{(a)}(\mathbf{x}')\right]$ (here $r(\mathbf{x}', \mathbf{x})$ is the probability transition density and $0 \le p^{(a)}(\mathbf{x}') \le 1$ defines the probability of a trajectory break; correspondingly, N is a random number of the break state). The random weights $\{Q^{(m)}\}$ from (1.10) are defined by the following recurrent ratios:

$$Q^{(0)} = \frac{f\left(\boldsymbol{\xi}^{(0)}\right)}{\pi\left(\boldsymbol{\xi}^{(0)}\right)}; \quad Q^{(m)} = Q^{(m-1)} \times \frac{k\left(\boldsymbol{\xi}^{(m-1)}, \boldsymbol{\xi}^{(m)}\right)}{p\left(\boldsymbol{\xi}^{(m-1)}, \boldsymbol{\xi}^{(m)}\right)}; \quad m = 1, ..., N. \tag{1.12}$$

Taking into account that the ratios (1.8) have the form (1.9), we get the following ***randomized projection-mesh functional algorithm***.

ALGORITHM 1. *Simulate n trajectories*

$$\boldsymbol{\xi}_j^{(0)}, \boldsymbol{\xi}_j^{(1)}, ..., \boldsymbol{\xi}_j^{(N_j)}; \quad j = 1, ..., n \tag{1.13}$$

of the applied Markov chain (1.11) *and get the values*

$$\tilde{\varphi}^{(\mathbf{x}_i)}(n) = \frac{1}{n} \sum_{j=1}^{n} \sum_{m=0}^{N_j} Q_j^{(m)} \kappa^{(\mathbf{x}_i)} \left(\boldsymbol{\xi}_j^{(m)} \right); \quad i = 1, ..., M;$$

here the wights $\left\{ Q_j^{(m)} \right\}$ *are calculated with respect to the formulas of the form* (1.12)*:*

$$Q_j^{(0)} = \frac{f\left(\boldsymbol{\xi}_j^{(0)}\right)}{\pi\left(\boldsymbol{\xi}_j^{(0)}\right)}; \quad Q_j^{(m)} = Q_j^{(m-1)} \times \frac{k\left(\boldsymbol{\xi}_j^{(m-1)}, \boldsymbol{\xi}_j^{(m)}\right)}{p\left(\boldsymbol{\xi}_j^{(m-1)}, \boldsymbol{\xi}_j^{(m)}\right)}; \quad j = 1, ..., n; \quad m = 1, ..., N_j.$$

Then approximate the function $\varphi(\mathbf{x})$ *with respect to the formula of the form* (1.5)*:*

$$\varphi(\mathbf{x}) \approx L^{(M)} \tilde{\varphi}(\mathbf{x}) = \sum_{i=1}^{M} w^{(i)} \left(\tilde{\varphi}^{(\mathbf{x}_1)}(n), ..., \tilde{\varphi}^{(\mathbf{x}_M)}(n) \right) \chi^{(i)}(\mathbf{x}). \tag{1.14}$$

In the works [2–4], the considerations of *the theory of conditional optimization* are presented. In particular, the expediency of using the "absolutely stable" finite functions of *the multi-linear approximation* (or Strang – Fix approximation [8] with the basis producing function $\beta^{(1)}(u)$, which is equal to the B-spline of the first order) on a regular mesh with the step h with respect to every coordinate

$$\chi^{(i)}(\mathbf{x}) = \beta^{(1)}\left(\frac{x^{(1)}}{h} - j_i^{(1)} \right) \times ... \times \beta^{(1)}\left(\frac{x^{(d)}}{h} - j_i^{(d)} \right); \tag{1.15}$$

$$\beta^{(1)}(u) = \begin{cases} u+1 \ \text{ for } \ -1 \le u \le 0; \\ -u+1 \ \text{ for } \ 0 \le u \le 1; \\ 0 \ \text{ otherwise;} \end{cases}$$

$$\mathbf{x} = \left(x^{(1)}, ..., x^{(d)}\right) \ \ \mathbf{x}_i = \left(j_i^{(1)} h, ..., j_i^{(d)} h\right); \ \ j_i^{(k)} \ \text{ are integer numbers;} \quad i = 1, ..., M \tag{1.16}$$

as basic functions (1.3) is proved; here the domain X, on which the solution $\varphi(\mathbf{x})$ of the equation (1.1) is approximated, is equal to cuboid. Moreover, it was proposed to use the "kernel" function from the ratios (1.7), (1.8) in the form

$$\kappa^{(\mathbf{x})}(\mathbf{y}) = \begin{cases} \frac{1}{h^d} \quad \text{for} \quad \mathbf{y} \in \Delta^{(\mathbf{x})}, \\ 0 \quad \text{otherwise,} \end{cases} \tag{1.17}$$

where $\Delta^{(\mathbf{x})} = \left\{ \mathbf{y} = \left(y^{(1)}, ..., y^{(d)}\right) : \ x^{(s)} - h/2 \le y^{(s)} \le x^{(s)} + h/2; \quad s = 1, ..., d; \right.$ $\left. \mathbf{x} = \left(x^{(1)}, ..., x^{(d)}\right) \right\}$.

For this case, the approximations of the coefficients (1.4) have the simplest form

$$w^{(i)}\left(\tilde{\varphi}^{(\mathbf{x}_1)}(n),...,\tilde{\varphi}^{(\mathbf{x}_M)}(n)\right)=\tilde{\varphi}^{(\mathbf{x}_i)}(n). \tag{1.18}$$

The Algorithm 1 with functions (1.15), (1.17) and approximation coefficients (1.18) is called in [2–4] as ***the multi-dimensional analogue of the polygon of frequencies method***.

In the Section 2 of this paper, we show that the described approach to construction of the projective-mesh Algorithm 1 is to a certain extent similar to construction of *the "kernel" estimators of probability densities* (see, for example, [9]). It is especially noted that in the the theory of "kernel" estimators (including [9]) the authors, who reason about density approximation, unduly not include the elements of the theory of numerical function approximation (see, for example, [5]). When adding this missing item, the corresponding "kernel" approximation of a probability density is essentially the same as the Algorithm 1.

2. Numerical approximation of probability densities using the "kernel" estimators

In the classical paper [9], the nonparametric estimator of a probability distribution density $f_{\hat{\boldsymbol{\xi}}}(\mathbf{x})$, $\mathbf{x}\in\mathbb{R}^d$ of the form

$$f_{\hat{\boldsymbol{\xi}}}(\mathbf{x})\approx Z_n(\mathbf{x})=\frac{1}{n}\sum_{j=1}^{n}\kappa^{(\mathbf{x})}\left(\hat{\boldsymbol{\xi}}_j\right), \tag{2.1}$$

using the sample values $\left\{\hat{\boldsymbol{\xi}}_1,...,\hat{\boldsymbol{\xi}}_n\right\}\subset\mathbb{R}^d$ from this distribution is considered. Here $\kappa^{(\mathbf{x})}(\mathbf{y})$ is some finite parametric, having the same shape for all values of the parameter $\mathbf{x}$ *"kernel" function*. The approximation (2.1) is called *the "kernel" estimator of the density* $f_{\hat{\boldsymbol{\xi}}}(\mathbf{x})$. By the way, in the paper [9] the term "kernel" is used without quotes. In this paper we use quotes in order to distinguish the names of the functions $k(\mathbf{x}',\mathbf{x})$ (this is the kernel of the integral equation (1.1)) and $\kappa^{(\mathbf{x})}(\mathbf{y})$.

For investigation of properties of the approximation (2.1), the following consequence of the large numbers law

$$Z_n(\mathbf{x})=\frac{1}{n}\sum_{j=1}^{n}\kappa^{(\mathbf{x})}\left(\hat{\boldsymbol{\xi}}_j\right)\approx\mathbf{E}\kappa^{(\mathbf{x})}\left(\hat{\boldsymbol{\xi}}\right)=\int\kappa^{(\mathbf{x})}(\mathbf{y})f_{\hat{\boldsymbol{\xi}}}(\mathbf{y})\,d\mathbf{y} \tag{2.2}$$

is used.

The evident constructive drawback of the "kernel" estimators theory (see, for example [9]) is related to absence of considerations on the algorithm for practical (firstly, numerical, computer) global approximation of the function $f_{\hat{\boldsymbol{\xi}}}(\mathbf{x})$ based on the theory of mesh function approximation (see, for example, [5]). Such an algorithm could look as follows.

Suppose that the random variable $\hat{\boldsymbol{\xi}}$ is distributed in the bounded domain $X \subset \mathbb{R}^d$ and consider the mesh (1.6) in this domain and also the approximation of the form (1.2) for the function $f_{\hat{\boldsymbol{\xi}}}(\mathbf{x})$:

$$f_{\hat{\boldsymbol{\xi}}}(\mathbf{x}) \approx L^{(M)} f_{\hat{\boldsymbol{\xi}}}(\mathbf{x}) = \sum_{i=1}^{M} w^{(i)} \left(f_{\hat{\boldsymbol{\xi}}}(\mathbf{x}_1), ..., f_{\hat{\boldsymbol{\xi}}}(\mathbf{x}_M) \right) \chi^{(i)}(\mathbf{x}).$$

ALGORITHM 2. *Calculate the values* $\tilde{f}_{\hat{\boldsymbol{\xi}}}^{(\mathbf{x}_i)}(n) = Z_n(\mathbf{x}_i)$; $i = 1, ..., M$ *with respect to the formulas of the form* (2.1) *and approximate the function* $f_{\hat{\boldsymbol{\xi}}}(\mathbf{x})$ *with respect to the formula of the form* (1.14)*:*

$$f_{\hat{\boldsymbol{\xi}}}(\mathbf{x}) \approx L^{(M)} \tilde{f}_{\hat{\boldsymbol{\xi}}}(\mathbf{x}) = \sum_{i=1}^{M} w^{(i)} \left(\tilde{f}_{\hat{\boldsymbol{\xi}}}^{(\mathbf{x}_1)}(n), ..., \tilde{f}_{\hat{\boldsymbol{\xi}}}^{(\mathbf{x}_M)}(n) \right) \chi^{(i)}(\mathbf{x}). \tag{2.3}$$

The Algorithm 2 is based on the analogs of the ratios (1.8):

$$\int f_{\hat{\boldsymbol{\xi}}}(\mathbf{y}) \kappa^{(\mathbf{x}_i)}(\mathbf{y})\, d\mathbf{y} \approx f_{\hat{\boldsymbol{\xi}}}(\mathbf{x}_i); \quad i = 1, ..., M, \tag{2.4}$$

which are in turn based on the ratios (2.1), (2.2).

Compare the Algorithms 1 and 2 and get the main conclusion of this paper.

REMARK 1. *The "kernel" Algorithm* 2 *for approximation of a probability density* $f_{\hat{\boldsymbol{\xi}}}(\mathbf{x})$*, based on approaches of the theory of mesh function approximation, is constructively equal to the randomized projection-mesh functional Algorithm* 1 *for approximation of the solution* $\varphi(\mathbf{x})$ *of Fredholm integral equation of the second kind* (1.1)*. Thus, it is expedient to use the new name* ***the randomized "kernel" functional algorithm*** *for the Algorithm* 1*.*

The only difference between the Algorithm 1 and 2 is defined by the distinction of forms of Monte Carlo estimators for approximate calculation of functionals (1.8) and (2.2) (which is related to the certain difference between functions $\varphi(\mathbf{x})$ and $f_{\hat{\boldsymbol{\xi}}}(\mathbf{x})$). The difference is also related to the fact that for the problem of approximation of the density $f_{\hat{\boldsymbol{\xi}}}(\mathbf{x})$, the sample $\left\{ \hat{\boldsymbol{\xi}}_1, ..., \hat{\boldsymbol{\xi}}_n \right\}$ is considered *to be given* (and the number n of sample values is fixed and cannot be increased), but for the function $\varphi(\mathbf{x})$ the number n of the simulated trajectories (1.13) of the applied Markov chain (1.11) *may vary.*

In connection with the main conclusion of the Remark 1, we can formulate the following considerations.

REMARK 2. *For the "kernel" Algorithm* 2 *for approximation of a probability density* $f_{\hat{\boldsymbol{\xi}}}(\mathbf{x})$ *we can use considerations of the theory of conditional optimization of the randomized projection-mesh functional Algorithm* 1 *from the works* [2–4].

By analogy of the works [2–4] we can recommend to use functions (1.15), (1.17), the mesh (1.16) and approximation coefficients of the form (1.18), i. e.

$$w^{(i)} \left(\tilde{f}_{\hat{\boldsymbol{\xi}}}^{(\mathbf{x}_1)}(n), ..., \tilde{f}_{\hat{\boldsymbol{\xi}}}^{(\mathbf{x}_M)}(n) \right) = \tilde{f}_{\hat{\boldsymbol{\xi}}}^{(\mathbf{x}_i)}(n), \tag{2.5}$$

in the Algorithm 2. In particular, it allows to get the following ratios for conditionally optimal parameters of the Algorithm 2:

$$M_{opt} = \left[\frac{\hat{H}_1[(2\nu+1)d+4]}{(2\nu+1)d}\right]^{d/2} \gamma^{-d/2}, \tag{2.6}$$

$$n_{opt} = \frac{\hat{H}_2^2 \hat{H}_1^{d/2}\,[(2\nu+1)d+4]^{2+d/2}}{16[(2\nu+1)d]^{d/2}} \times (2\ln M_{opt} - \ln\ln M_{opt} + \hat{H}_3) \times \gamma^{-2-d/2} \tag{2.7}$$

for the fixed error level $\gamma > 0$ and the specially selected positive constants $\hat{H}_1$, $\hat{H}_2$, $\hat{H}_3$ and ν.

Taking into account the fact that the basis (1.15) is "modelled", it is possible to recommend to use the normalized function $L^{(M)}\tilde{f}_{\hat{\boldsymbol{\xi}}}(\mathbf{x})$ from the ratio (2.3) (with functions (1.15), (1.17) and coefficients (2.5)) as a density for numerical simulation of additional sample values $\{\boldsymbol{\xi}_j\}$, close in distribution to the values $\left\{\hat{\boldsymbol{\xi}}_j\right\}$ from the ratio (2.1), using the corresponding version of the discrete superposition method (see Sections 17, 18 of the book [10]).

REMARK 3. *For development of the theory of construction and conditional optimization of the randomized projection-mesh functional Algorithm* 1 *it is possible to use the considerations of the theory of "kernel" estimators of probability densities from the paper* [9].

Conclusion

In this paper, it is shown that if include the theory of numerical function approximation, then it is possible to get the "constructive" (practical) version of the "kernel" estimator for approximation of an unknown probability density $f_{\hat{\boldsymbol{\xi}}}(\mathbf{x})$ using the given sample values (Algorithm 2). This construction is analogous to the randomized projection-mesh functional algorithm for approximation of the solution $\varphi(\mathbf{x})$ of Fredholm integral equation of the second kind (1.1) (i. e. to the Algorithm 1). If additionally choose the approximation basis (1.15), the uniform mesh (1.16) in the cuboid X, the "kernel" function (1.17) and the coefficients (2.5) (i. e. to consider the version of the Algorithm 2, which corresponds to the multi-dimensional analogue of the polygon of frequencies method) and use the technique presented in the works [2–4], then we can get the ratios (2.6), (2.7) for conditionally optimal parameters of the Algorithm 2.

References

[1] Mikhailov G.A. (2000). *Weighted Monte Carlo Methods*. SD RAS Publisher, Novosibirsk [In Russian].

[2] Voytishek A.V. (2001). *Discrete-Stochastic Numerical Methods / Doctorial dissertation*. Novosibirsk [In Russian].

[3] Mikhailov G.A., Voytishek A.V. (2006). *Numerical Statistical Modelling. Monte Carlo Methods*. Publishing House "Akademia", Moscow [In Russian].

[4] Voytishek A.V. (2007). *Functional estimators of the Monte Carlo Method*. NSU Publisher, Novosibirsk [In Russian].

[5] Bahvalov N.S. (1975). *Numerical Methods*. Nauka, Moscow [In Russian].

[6] Voytishek A.V., Shipilov N.M. (2017). On Randomized Algorithms for Numerical Solution of Applied Fredholm Integral Equations of the Second Kind *AIP Conference Proceedings*. **1907**, 030015.

[7] Voytishek A.V. (2018). Development and optimization of randomized functional numerical methods for solving the practically significant Fredholm integral equations of the second kind. *Journal of Applied and Industrial Mathematics*. Vol. **12**, No. 2. pp. 382–394.

[8] Marchuk G.I., Agashkov V.I. (1981). *Introduction to Projection–Mesh Methods*. Nauka, Moscow [In Russian].

[9] Epanechnikov V.A. (1969). Nonparametric estimation of a multidimensional probability density. *Theory of probability and its Applications*. Vol. **14**, No. 1. pp. 153–158.

[10] Voytishek A.V. (2007). *Edditional Information about Numerical Modelling of Random Elements*. NSU Publisher, Novosibirsk [In Russian].

Investigation of the chi-squared test errors

Oleg A. Makhotkin
The Institute of Computational Mathematics and Mathematical Geophysics, Novosibirsk, Russia,
Novosibirsk State University, Novosibirsk, Russia
e-mail: oam@osmf.sscc.ru

Abstract

In this paper, the second kind errors, which arise in the problem of testing the hypothesis about the form of the sample density, are considered. The class of the errors associated with the use in the test statistic the piecewise constant approximation instead of the original density, was investigated in detail. The results of the theoretical analysis and computer experiments are presented and discussed.
Keywords: Chi-squared test, testing of hypothesis, L_1 distance,errors of the second kind

Introduction

The chi-squared statistic is frequently used for testing the hypothesis about the form of the sample probability distribution density function(p.d.d.f.). The null hypothesis is simple. It states: the sample p.d.d.f $p_s(x) = p(x)$. The alternative hypothesis is complex. It states: $p_s \neq p(x)$. The test has the four results: two correct and two incorrect. The incorrect results are: the null hypothesis H_0 is true, but it is rejected, the H_0 is false, but it is adopted.The value of the error of the I kind is defined by the researcher(the confidence level α). The value of the error of the II kind can be estimated only for a concrete probability density using the statistical simulation method. In this paper, the alternative approach is investigated. It uses the L_1 -distance between the tested p.d.d.f.[1] $p(x)$ and the alternative densities p_{alt}.

1 The structure of the chi-squared test

Let us suppose(for simplicity), that a tested p.d.d.f. $p(x)$ is defined on $(0, 1)$. The interval $(0, 1)$ can be represented as an union of m subintervals $\Delta_i = \{x : x_{i-1} < x < x_i\}_{i=1}^m$, where x_i are the knots of the grid $X_m = \{x_0 = 0 < x_1 < \ldots < x_m = 1\}$. The piecewise constant function($h_i = x_i - x_{i-1}$)

$$p_{PC}(x) = \sum_{i=1}^{m} \frac{P_i}{h_i} \chi(x|\Delta_i) = \sum_{i=1}^{m} c_i \chi(x|\Delta_i) \tag{1}$$

[1] probability distribution density function(s)

is the Galerkin approximation of p.d.d.f. $p(x)$. Here the functions $\chi(x|\Delta_i) = 1$, if $x \in \Delta_i$, and $= 0$ otherwise, $P_i = \int_0^1 p(x)\chi(x|\Delta_i)dx$.
A sample $\mathcal{S}_N = \{\eta_1, \ldots, \eta_N\}$ defines the random piecewise density

$$\bar{p}_{PC}(x) = \sum_{i=1}^{m} \frac{m_i}{Nh_i}\chi(x|\Delta_i) = \sum_{i=1}^{m} \frac{B_i}{h_i}\chi(x|\Delta_i) = \sum_{i=1}^{m} \bar{c}_i\chi(x|\Delta_i), \tag{2}$$

where $m_i = \sum_{j=1}^{N} \chi(\eta_j|\Delta_i)$. According to the theory, statistics

$$\chi^2 = \sum_{i=1}^{m} \frac{(N \cdot P_i - m_i)^2}{N \cdot P_i} = N\sum_{i=1}^{m} \frac{(P_i - B_i)^2}{P_i} = N\sum_{i=1}^{m} h_i\frac{(c_i - \bar{c}_i)^2}{c_i} \tag{3}$$

asymptotically has the chi-squared distribution with $m-1$ degrees of freedom, The hypothesis $H_0: \ p_{sample}(x) = p(x)$ is adopted if $\chi^2 < \chi^2_{m-1,\alpha}$, $\chi^2_{m-1,\alpha}$ is the upper critical value for confidence level α ($P(\chi^2_{m-1} > \chi^2_{m-1,\alpha}) = \alpha$).

2 Error of approximation

The chi-squared statistics (3) includes the piecewise constant approximation $p_{PC}(x)$ instead of the test density $p(x)$. The natural measure of the distance between any two p.d.d.f. $p(x), g(x)$, defined on $X = (0,1)$, is the L_1-norm of their difference

$$L(p,g) = ||p(x) - g(x)|| = \int_X |p(x) - g(x)|dx.$$

According to the Scheffe theorem [1],

$$L = 2\sup_B |\int_B p(x)dx - \int_B g(x)dx| = 2\sup_B |P(B) - G(B)|,$$

where $\{B\}$ are the Borel sets of X.

The value of $L(p, p_{PC})$ defines the error of approximation of p.d.d.f. $p(x)$ by the piecewise p.d.d.f. $p_{PC}(x)$. For the calculation, it can be represented in the form

$$L(p, p_{PC}) = h_t\sum_{i=1}^{m} h_i\sum_{j=1}^{MG}\int_0^1 f(x_{i-1} + h_ih_t(j-1+t))dt,$$

where $h_t = 1/MG$. $f(x) = |p(x) - p_{PC}(x)|$. The integrals over t can be approximated by the Gauss quadrature with two knots. Successively increasing MG, it is possible to obtain the integral estimation with the desirable number of the correct digits.
Example 1: the probability density

$$p(x) = \begin{cases} \frac{27}{2}x^2, & 0 \le x \le \frac{1}{3}, \\ \frac{27}{2}[x^2 - 3(x-\frac{1}{3})^2], & \frac{1}{3} \le x \le \frac{2}{3}, \\ \frac{27}{2}(1-x)^2, & \frac{2}{3} \le x \le 1 \end{cases} \tag{4}$$

was approximated on the non-uniform grid with m intervals. The knots of the grid were calculated by the formula: $x_0 = 0$, $x_1 = 2/(m+2)$, $x_i = x_1 + (i-1)/(m+2)$, $i = 2..m-1$, $x_m = 1$. The calculated values of distances equal L(m=10)=1.105e-1, L(m=20)=5.399e-2.

3 The artificial test densities

Formula (3) shows, that the chi-squared statistic contains the piecewise approximations of the p.d.d.f. $p(x)$. The natural question can be formulated as follows: "Are there the probability densities with the same values of $\{c_i\}$ as the test p.d.d.f. $p(x)$?". The p.d.d.f. $f(x)$ gives the answer to the formulated question. It has the form

$$f(x) = \sum_{i=1}^{m} c_i g_i(t_i(x))\chi(x|\Delta_i) = \sum_{i=1}^{m} c_i(1 + \phi(t_i(x))\chi(x|\Delta_i),$$

where $t_i(x) = (x - x_{i-1})/h_i$, $g(t)$ is an arbitrary p.d.d.f. on (0,1), $\phi(t) = g(t) - 1$. The distance $L(f, p_{PC}) = \int_0^1 |\phi(t)|dt$. The formula for p.d.d.f. $f(x)$ shows that there are uncountable many probability densities with the same piecewise constant approximation as the tested p.d.d.f. $p(x)$.
The next two examples show the two classes of the alternative densities for which we obtain the same value of χ^2 as for $p(x)$.
Example 2: $g(t)(t) = 2sin^2(\pi t)$, $L = 2/\pi = 0.636$. For this $g(t)$ p.d.d.f. $f(x)$ is continuous with the first derivative.
Example 3: the interval $(0, 1)$ is divided into n equal parts. In $(n - 1)$ subintervals $g(t) = 1 - r$, $0 < r < 1$, in the last one $g(t) = 1 + (n - 1)r$. $L = 2r - 2r/n$. For $r - 0.5$, $n = 5$ L is equal to 0.8. Fig. 1 shows the p.d.d.f. $f(x)$ for Example 2.

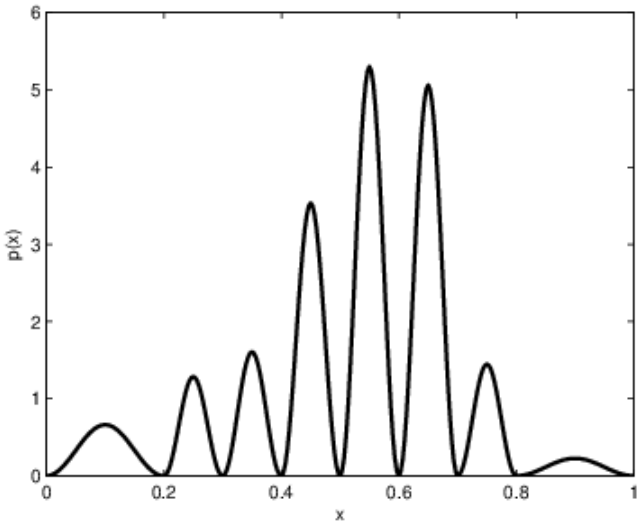


Fig. 1. The probability density $f(x)$ with $g(t) = 2\sin^2(\pi t)$.

Adopting for the use the test density $p(x)$ instead the $f(x)$ from the example 2 causes sufficiently large differences in the probability of events: $\Delta P_{max} = 0.318$. However adopting the principle of non-complexity to the practical distributions removes these artificial densities. As the measure of complexity of the p.d.d.f. its full variation can be used. For the p.d.d.f. in Fig. 1 $\bigvee(f(x)) = 38.22$, the paternal density has $\bigvee(p(x)) = 4.5$.

4 The piecewise constant alternative densities

The chi-squared statistics χ^2 can be represented as the function of two probability distributions $P = \{P_i\}$ and $B = \{B_i = m_i/N\}$:

$$\chi^2(P, B) = N \sum_{i=1}^{m} (P_i - B_i)^2 / P_i.$$

For finding the probability distribution Y, the most deviating from P, the following conditional maximum problem must be solved:

$$\Psi(Y, P) = \sum_{i=1}^{m} |P_i - Y_i| \ \rightarrow \ \max_{Y},$$

$$Y_i \ \geq \ 0, \ \sum_{i=1}^{m} Y_i = 1,$$

$$\chi^2(Y, B) = N \sum_{i=1}^{m} \frac{(Y_i - B_i)^2}{Y_i} \ \leq \ \bar{\chi}^2.$$

Example 4: p.d.d.f. (4) was used for calculations. It was approximated on the grid $X_{10} = \{x_i = i/10\}_{i=0}^{10}$. The theoretical probability distribution P is presented in the first column of the Table 1. Its sample estimation B with $N = 100$ equals $0, 0.05, 0.08, 0.1, 0.22, 0.22, 0.21, 0.09, 0.033, 0$. The chi-squared statistic $\chi^2(P, B) = 5.86$, Upper critical value is equal to $\chi^2_{7,0.05} = 16.9$. The results of optimization is presented in Table 1. The theoretical distribution P was used as the initial state vector for the optimization procedure.

Tab. 1. Optimal vectors Y for two values of $\bar{\chi}^2$.

i	P_i	$Y_i(\bar{\chi}^2 = 4)$	$Y_i(\bar{\chi}^2 = 5.86)$
1	4.50e-3	4.89e-4	2.29e-4
2	3.15e-2	6.67e-2	6.66e-2
3	8.55e-2	6.67e-2	1.20e-1
4	1.63e-1	9.04e-2	8.08e-2
5	2.16e-1	2.82e-1	1.68e-1
6	2.16e-1	1.88e-1	2.52e-1
7	1.63e-1	2.14e-1	2.27e-1
8	8.55e-2	7.43e-2	6.64e-2
9	3.15e-2	1.75e-2	1.91e-2
10	4.50e-3	3.91e-4	2.29e-4
		$\Psi = 0.30$	$\Psi = 0.34$

For $Y_i(\bar{\chi}^2 = 10)$ the value $\Psi = 0.35$ was obtained.
The Fig. 2 shows the piecewise constant densities for the probability distribution P and for the optimal probability distribution Y.

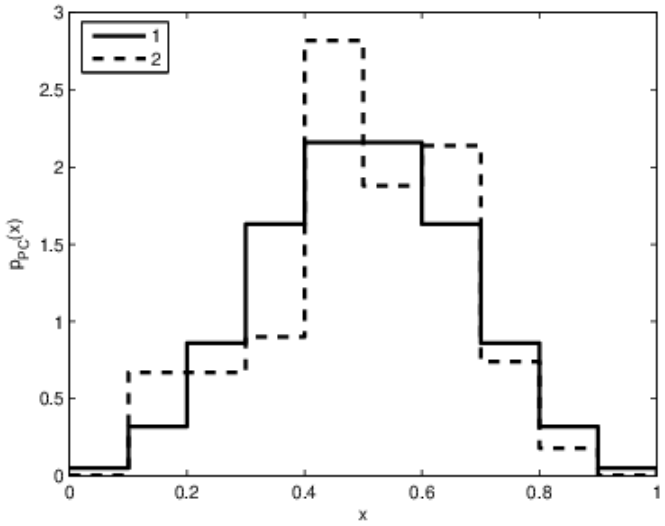


Fig. 2. The piecewise probability densities: 1 - for tested distribution P, 2 - for the optimal distribution $Y_(\bar{\chi}^2 = 4)$.

Summary

In this paper, the computer procedures were used for the calculations the errors, associated with the use in the chi-squared test the piecewise constant approximations instead of the real densities. Examples 2,3 show that there exist the uncountable many artificial densities which have the same piecewise constant approximation as the paternal density. However adopting the principle of non -complexity of the practical densities avoids these cases. This principle demands that the practical densities were not very complex. As the measure of complexity, the full variation of the density can be used.

References

[1] Scheffe H., A useful convergence theorem for probability distributions, Annals of Math. Statistics, vol. 18, 1947, pp.434-438.

Comparison of generalisation error's methods on case of Linear Regression

Pavel Peresunko[1],Kristina Pakhomova[2], Eugenia Soroka[3] and Sergey Videnin[4]
Institute of Space and Information Technology, Sibirian Federal University. Russia
e-mail: [1]peres94@yandex.ru,[2]kpakhomova@sfu-kras.ru,[3]jollot@yandex.ru,
[4]svidenin@sfu-kras.ru

Abstract

In feature selection process it is important to estimate generalized error of selected model. There are a lot of methods based on resampling procedure. All this methods have their features, and this work have a large study. Each of selected method was run for 7200 times to find any pattern according to samples and feature count of dataset. This research shows accuracy of each method for Lasso regression in case of linear function.

Keywords: Generalization Error Estimation, Cross-Validation, Monte Carlo Cross Validation, Cross-validation and cross-testing, Nested Cross Validation, Lasso

Introduction

The goal of prediction regression analysis is to find a model that describes dependencies in an accurate way. That model should have the small expected generalization error. It was proven that the expected generalization error of the model decomposes to an irreducible error, a bias and a variance [1]. The model may have hyperparameters, which control the learning process. Hyperparameters tuning may increase the bias and decrease the variance, and vice versa. Hyperparameters' values can be selected via a grid search [2]. For example, in [3] a penalty term λ was selected by the grid search for Lasso regression. Model's coefficients of Lasso regression depend on the value of the penalty term λ and each unique value of λ may create a new model. The desired model may be selected with model selection procedures. There are several alternatives in the literature about the estimation of the expected generalized error of the selected model.

For example, available samples may be divided into two parts: train and test [4]. To measure generalized error, the model is fitted on train samples and makes a prediction for unseen test samples. Then error on test samples is computed. RMSE is one of the examples of error's measure. Generalization error of the model may be estimated with this approach if the model has no hyperparameters. If the model has hyperparameters and they are chosen as based on minimization of test error, test error will be biased and underestimate true generalization error. To overcome it, all available samples are split into three parts: train, validation and test [5]. The model is fitted on train samples. Models are selected on the validation part. With grid search, sets of hyperparameters are created. Error on validation samples

are evaluated for each set of hyperparameters and hyperparameters' values with the smallest error are selected. The model with selected hyperparameters is refitted on train and validation samples, and then test error is evaluated on test samples. This test error approximates the generalization error.

Train and test split works well with a large number of samples. If the size of the dataset is small, test error will depend on how exactly the dataset was split. Monte Carlo cross validation may be used to reduce dependency on how the dataset was split into train and test sets [13]. In this approach, all available samples split at random on train and test samples N times. For each of N splits model is fitted on train samples and error is evaluated on test samples. After that, generalization error is estimated by the average of N test errors. It has been proven that this method is asymptotically consistent [7].

If the model has hyperparameters, all available samples should be separated into train, validation and test samples. In Monte Carlo cross test validation approach, this random separation on three parts should be repeated N times. The model is fitted on train samples; model's hyperparameters are tuning to have the smallest error on validation samples. Then test error is evaluated on test samples for each separation. Average of N test errors is the estimation of generalization error. Another solution to the lack of samples is ?ross Validation (CV). In cross validation, all available samples are split into k blocks. One of the blocks is used as a test set and other blocks are used as a training set. Each of k blocks is selected as left out samples and the test error is evaluated on each selected blocks in the loop. The average of test errors can be considered as an estimate of the true generalization error. It was shown that the CV error is an almost unbiased estimate of generalization error [8].

The problem of CV is the separation step. CV split samples into train and test samples, so it works well if the model does not have hyperparameters. If the model has hyperparameters and they are selected based on CV error, the CV error may be biased and CV will underestimate the true generalization error. In that case, Nested CV can be used [8]. As in CV, in Nested CV all samples split into k blocks. On the next step, one block is used as a test set. The rest of $k-1$ training blocks are used in a CV to get a model's hyperparameters with the loser CV error. Then the model is refitted on $k-1$ training blocks with selected hyperparameters, and error on the test set is evaluated. Then next block is used as a test, and the rest of the blocks as a training set and so on. In the results, k errors are evaluated. The average of these errors is the estimation of the generalization error.

While Nested CV error has high accuracy, it has some disadvantages. Firstly, it does not show which exactly hyperparameters are better on unseen data [9]. Secondly, different models are selected on each of k blocks, and it does not allow to investigate the final model. To overcome these drawbacks Cross validation and cross-testing method was introduced [9]. In this approach, all available samples are divided into train and test samples. Model's hyperparameters with smallest cross validation error on train data are chosen. Then test samples are divided into m parts. One of the m test parts is selected and added to train dataset. Other parts are used to evaluate test error. On the next iteration, another part from m test parts are selected and added

to original train dataset, and the test error is evaluated on the remaining part. This process continues until each of m test blocks was added to train samples at least once. The average test error is the estimation of the true generalization error. It has been shown [9] that sometimes CV and cross testing outperform nested cross validation.

1 Experiment

Methods of generalization error estimation are compared on real datasets in some researches [10],[11]. The problem of these experiments is an unknown irreducible error in data and real generalization error of the model. For this reason, simulation data was created to test the generalization error estimate.

The main goal in each experiment is built Lasso regression and estimate its generalization error. There are two reasons to use. Firstly, it builds a linear model with hyperparameters value, so methods of hyperparameters selection can be used in these settings. Secondly, it is fast to fit this model.

Assume that the model has an input vector $X^T = (X_1, X_2, ..., X_m)$ and wants to predict real valued Y. Suppose that true dthe ata generating process comes from linear equitation

where ε is random error, β_j is slope coefficients, X_j is random variable. Mean value of ε is 0, standard deviation of ε equal to σ and generated from a Gaussian distribution, i.e. ε $N(0, \sigma_\varepsilon^2)$. Random error ε is independent of X.

Suppose that each of random variables X_j comes from Gaussian distribution with mean is 0 and standard equitation is 1, i.e. X_j $N(0, 1)$. With those assumptions, there is no reason to center and scale random values, or model's input vector.

The first half of the coefficients β_j is equal to 0, so they are not significant. Count of sithe gnificant coefficient is $m^* = [m/2]$. With these settings, Lasso should remove input variables with the coefficients equal to 0.

For each significant variables, the value of its coefficient β_j is given by

$$\beta_j = \frac{\sqrt{j}}{\sqrt{\sum_{k=1}^{m^*} k}}$$

Through this approach, all significant coefficients form an arithmetic progression and variance of Y do not depend on the count of random variables. Calculate the variance of Y to prove it.

$$Var(Y) = Var(\sum_{j=1}^{m} X_j\beta_i + \varepsilon) = \sum_{j=1}^{m} \beta_j^2 Var(X_j) + Var(\varepsilon) =$$
$$= \sum_{j=1}^{m^*} \beta_j^2\sigma_x^2 + \sum_{j=m^*+1}^{m} 0 * \sigma_x^2 + \sigma_\varepsilon^2 = \sigma_x^2 \sum_{j=1}^{m^*} \frac{\sqrt{j}^2}{(\sqrt{\sum_{k=1}^{m^*} k})^2} + \sigma_\varepsilon^2 =$$
$$= \sigma_x^2 \frac{\sum_{j=1}^{m^*} j}{\sum_{j=1}^{m^*} k} + \sigma_\varepsilon^2 = \sigma_x^2 + \sigma_\varepsilon^2$$

Therefore, the variance of Y depends on the variance of a random variable and the variance of random error. On the first step of the proof, the variance of two independent random variables is equals to the sum of the variance of each variable. On the second step, the random variable scaled by a constant, and the variance is scaled by the square of that constant.

Calculate the irreducible error of linear regression for this task. Assume that the coefficient of linear regression is equivalent to the real coefficient of random variables, i.e. $\hat{\beta}_j = \beta_j$, $j = 1, 2, ..., m$. In that case, irrean an ducible error can be derived by

$$E\{(E(Y|X_1, X_2, ..., X_m) - y)^2\} = E\{(\sum_{j=1}^{m} X_j\beta_j + \varepsilon - \sum_{j=1}^{m} X_j\hat{\beta}_j)^2\} = E\{(\varepsilon)^2\} =$$
$$= E\{(0 - \varepsilon)^2\} = E\{(M\{\varepsilon\} - \varepsilon)^2\} = Var\{\varepsilon\} = \sigma_\varepsilon^2$$

Experiment's parameters It is shown that the irreducible error in this task equals the variance of random error. To test the estimation of generalizing error by different methods in different situations, count of samples n and count of random variables m should vary from experiment to experiment. Assume that values of n are generated by this rule:

$$N = \{x | x \in [1/42^{i+1} + 2i + 1], i = \overline{2, 17}\}$$

where [.] is the operation of rounding argument to the nearest whole number.

With this rule, set values of n is $N \sim \{8, 11, 15, 19, 25, 32, 40, 52, 68, 90, 120, 163, 221, 304, 421, 586\}$. It makes sense to use exponential low, as qthe uality of error estimation grows with ca ount of samples. We can use inverse transformation to get transformed back values of sample sizes, $N^{'}$ $\{3, ..., 20\}$. It can be used in futhe ture to plot the results of methods.

Suppose that values of m are generated by this rule:

$$M = \{x | x \in [1/36^{i+1} + 0.13], i = \overline{0, 14}\}$$

where [.] is the operation of rounding argument to the nearest whole number.

With this rule, set values of m is
$M \sim \{1, 2, 3, 4, 5, 6, 9, 12, 16, 22, 30, 40, 55, 74, 101\}$. Model's input count often is less than count of samples.

Model building and error estimation are made for each combination of n and m values in this experiment. For each combination of samples' count and input's count, error estimation depends on which exactly samples are generated. To reduce this dependency, error estimation was made for 30 times for each combination of n and m. Different methods of error estimation were performed on the same set of samples to hold them under the same conditions.

For each set of n and m test samples was created too. Count of test samples is 100000. With this count, test samples are a good estimation of the true generalization error of the model by the law of large numbers. Test error is the average of an error on each test sample, so it will almost surely converge to mean error.

True mean μ is linear combination of inputs and hence it is under Gaussian distribution. It was shown that $\mu = E(Y|X_1, X_2, ?, X_m)$ have variance σ_X^2. If $\sigma_X^2 = 1$, then μ if lied from -2 to 2 with probability 95.44% by three-sigma rule. Then set $\sigma_\varepsilon = 0.4$ to random error lies from -0.8 to 0.8 with probability of 95.44%.

For CV of count blocks $k = 10$ [12]. For Monte Carlo Cross Validation 70 % and 80% samples were selected as train samples [13]. Count of splitting was 10. In another research, this count should be greater, but experiments will take too a lot of time to run. For Monte Carlo Cross test validation, validation's part is 20% and the test part is 30%. In another experiment, validation's part is 15%, the test is 20%.

2 The results of Experiment

Table with error estimation was created as the results of this experiment. For each combination of n, m and method of error estimation 30 error estimations for different datasets are created. Also for all rows, the absolute difference between the estimation of the generalization error and true error is evaluated. This value is the error of the error estimation method. This table has 57600 rows. Mood's median test [14] was performed to check if all methods have the same error. As a result, the $p - value$ is lower than 0.05, and to check which one methods differ, the post-hoc test is used. For all pairs of methods Mood's median test was conducted. Adjustments to the p-values were made to avoid inflating the possibility of making a type-1 error. Benjamini and Yekutieli correction was used [15]. Table 1 gives the results of these tests. Methods with the same latter have an insignificant difference in median's value. The threshold for $p - value$ is 0.05.

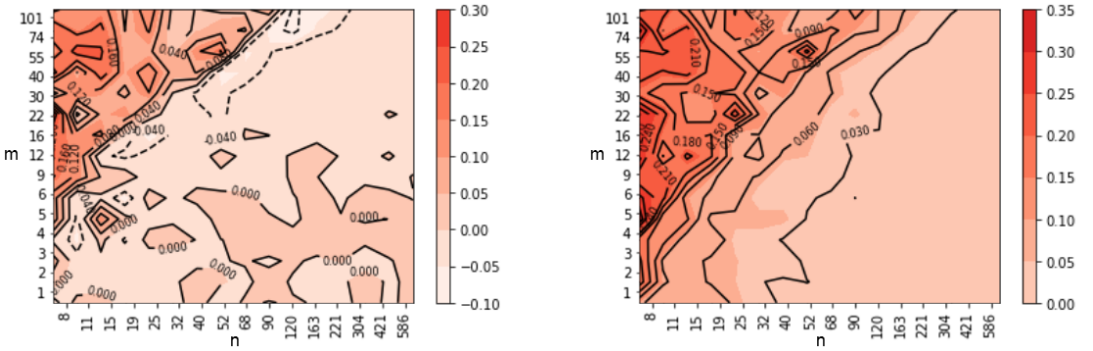
Methods with a lower median of error are a CV, MCCV, and Nested CV, their median does not differ significantly.

For all combinations of samples count n and feature count m differences in errors' median are not significant. But it is important to find out the best method for different m and n. Figure 3 contour plot of the median absolute difference between true errors and estimated by MCCV errors are shows 3. If the count of samples is

Table 1: Pairwise median's test results

?	Method	Parameters	Latter	Error's median
1	CV	k = 10	A	0.035
2	Nested CV	k = 10	A	0.036
3	CVCT	70%	B	0.059
4	CVCT	80%	C	0.066
5	MCCV	70%	D	0.04
6	MCCV	80%	A	0.035
7	MCCTV	20% and 30%	E	0.043
8	MCCTV	15% and 20%	DE	0.042

greater than the features' count, a median of error is small and not differ a lot. In another case the difference of errors are big. In that area all methods underestimate true error, so makes sense to divide these areas and work on each of them separately.



a) Contour plot of median error b) Contour plot of median absolute error

Figure 1: The dependence between methods

Figure 2 for each n and m number of best method with minimal error is shown. In a situation, where samples' size is lower than features count, 35 times out of 48 MCCV has the smallest error. It means that in most cases MCCV is the best method for error estimating if the count of samples is lower than the count of features. If $n > m$, Cross validation 86 times out of 181 times has the smallest absolute error, Nested Cross Validation 62 times out of 181 has the smallest absolute error. But there is no pattern of how to choose the method as they are scattered uniformly.

3 Discussion

A different method of true error estimation is considered in this work. It is shown that in the case of linear model best methods are a CV, Nested CV, and MCCV. If the count of samples is lower then count of features, it ought to use MCCV. In

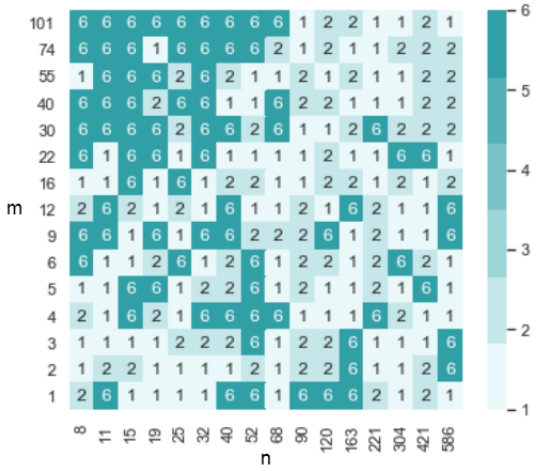


Figure 2: Number of the method with minimum absolute error

another case, CV or Nested CV can be used. It is ought to note that Nested CV works $k-1$ times longer more than a CV. For Monte Carlo, Cross Validation count of repition should be greater, so maybe it's quality will be better.

It is important to remember that in these experiments linear dipendency was used. The Lasso have linear structure too, so it can approximate the true functioan connection. Usually, the researcher does not know real structre of the model.

Lasso model is not flexible, so maybe that is the reason why CV and MCCV work better than a nested CV. In the next researchers, another dataset and more flexible models are planned to be used.

References

[1] Hastie T., Tibshirani R., Friedman, J. (2001). *The Elements of Statistical Learning*. New York, NY, USA: Springer New York Inc.

[2] Kazius J., Mcguire R., and Bursi R. (2005). Derivation and Validation of Toxicophores for Mutagenicity Prediction. *McGuire, Bursi*. pp. 312-320.

[3] Tibshirani R. (1996). Regression Shrinkage and Selection via the lasso. *Journal of the Royal Statistical Society*. pp. 267-288.

[4] Alpaydin E. (2014). *Introduction to machine learning*. MIT press.

[5] Hastie T, Tibshirani R, Friedman J, Franklin J. (2005). The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*. pp. 83-85.

[6] Xu Q. S., Liang Y. Z., and Du Y. P. (2004). Monte Carlo cross-validation for selecting a model and estimating the prediction error in multivariate calibration. *Journal of Chemometrics*. pp. 112-120.

[7] Shao J. (1993). Linear Model Selection by Cross-Validation on JSTOR. *American Statistical Association*. pp. 486-494.

[8] Varma S, Simon R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* .

[9] Korjus K., Hebart M. N., and Vicente R (2016). An efficient data partitioning to improve classification performance while keeping parameters interpretable. *PLoS ONE*.

[10] Krstajic D., Buturovic L. J., Leahy D. E., and Thomas S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of Cheminformatics*. pp. 1-15.

[11] Kohavi R (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *Appears in the International Joint Conference on Articial Intelligence (IJCAI)*. pp. 1-7.

[12] McNeish D. M. (2015). Using Lasso for Predictor Selection and to Assuage Overfitting: A Method Long Overlooked in Behavioral Sciences. *Multivariate Behavioral Research*. pp. 471-484.

[13] Xu Q. S., Liang Y. Z., and Du Y. P. (2004). Monte Carlo cross-validation for selecting a model and estimating the prediction error in multivariate calibration. *Journal of Chemometrics*. pp. 112-120.

[14] Brown G. W., Mood A. M. (1951). On Median Tests for Linear Hypotheses. Proceedings of the Second Berkeley Symposium on Mathematical Statististics and Probility. *University of California Press*. pp. 159-166.

[15] Benjamini Y., and Yekutieli D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*. pp. 1165-1188.

On the distribution of the $MIN3$ two-sample test statistic

Petr Philonenko and Sergey Postovalov
Novosibirsk State Technical University, Novosibirsk, Russia
e-mail: petr-filonenko@mail.ru, postovalov@ngs.ru

Abstract

In our earlier papers, the $MIN3$ nonparametric two-sample test has been proposed. The statistic of the $MIN3$ test is a minimum of three dependent random variables. These random variables are p-values of the weighted Kaplan-Meier test, Bagdonavičus-Nikulin test based on the MCE model and Bagdonavičus-Nikulin test based on the SCE model.

The asymptotical distribution of the $MIN3$ two-sample test statistic is unknown. However using Monte-Carlo simulation, we found that the distribution of the $MIN3$ test statistic can be approximated by the Beta distribution of the third kind.

In the paper, we study the distribution of the $MIN3$ test statistic by the Monte-Carlo simulation. Estimated parameters and lower percentage points of the $MIN3$ test statistic distribution are represented in the paper.

Keywords: survival analysis, randomly right-censored observations, hypothesis testing, two-sample problem, Monte-Carlo simulation, $MIN3$ test, Bagdonavičus-Nikulin SCE test, Bagdonavičus-Nikulin MCE test, Weighted Kaplan-Meier test.

Introduction

One of the steps of hypothesis testing [1] consists in a comparison of the value of a test statistic with some critical value that depends on the corresponding distribution of the test statistics. It is known that the distribution of the test statistic under limited samples size may differ from the corresponding limit distribution [2]. In tasks with randomly right-censored observations, the distribution of censored time $F^C(t)$ and censoring rate r can affect the distribution of the test statistic.

In our earlier papers, we propose the $MIN3$ test [3] for two-sample problem. The asymptotical distribution of the $MIN3$ test statistic is unknown. This makes it difficult to apply the $MIN3$ test in practice because every time it is required to use special methods for the simulation of the test statistic distribution. For example, this problem can be solved by the Monte-Carlo method [4]. In this paper, we study the distribution of the $MIN3$ test statistic varying samples size, distribution of censored time, and censoring rate by the Monte-Carlo simulation. According to the results of the simulation, we estimate the parameters of the distribution of the $MIN3$ test statistic using the Beta distribution of the third kind.

This research has been supported by the Russian Ministry of Education and Science as a part of the state task (application number 1.1009.2017/4.6).

In Section 1, we consider the problem statement and describe the model of randomly right-censored observations. In section 2, we present the statistics of the $MIN3$ test. In Section 3, we study the distribution of the $MIN3$ test statistic and present the results of a Monte Carlo simulation.

1 Problem Statement

Suppose that we have two samples of continues variables ξ_1 and ξ_2 respectively, $X_1 = \{t_{11}, t_{12}, ..., t_{1n_1}\}$ and $X_2 = \{t_{21}, t_{22}, ..., t_{2n_2}\}$ of two survival distributions $S_1(t)$ and $S_2(t)$. The samples size are n_1, n_2 (if $n_1 = n_2$, then may be denoted as n). The observation $t_{ij} = \min(T_{ij}, C_{ij})$, where T_{ij} and C_{ij} are the failure and censoring times for the j-th object of the i-th group. T_{ij} and C_{ij} are i.i.d. with a cumulative distribution function (CDF) $F_i(t)$ and $F_i^C(t)$ respectively. Survival curve means the probability of survival in the time interval $(0, t)$

$$S_i(t) = P\{\xi_i > t\} = 1 - F_i(t).$$

Then the null hypothesis is

$$H_0 : S_1(t) = S_2(t)$$

against alternative hypothesis

$$H_1 : S_1(t) \neq S_2(t).$$

Let an indicator of censoring c_{ij} be equal 0 if t_{ij} is a failure time and be equal 1 if t_{ij} is a censored time.

Further, we consider a test statistic of the $MIN3$ test.

2 $MIN3$ Two-sample Test

A test statistic of the $MIN3$ test [3] is

$$S_{MIN3} = \min\{p_{WKM}, p_{BN2}, p_{BN3}\},$$

where

$$p_{WKM} = 2 \cdot \min\left\{F_{N(0,1)}(S_{WKM}), 1 - F_{N(0,1)}(S_{WKM})\right\},$$
$$p_{BN2} = 1 - F_{\chi^2(3)}(S_{BN2}),\ p_{BN3} = 1 - F_{\chi^2(2)}(S_{BN3}),$$

S_{WKM} is a test statistic of the Weighted Kaplan-Meier test (the test statistic represented in [5]), S_{BN2} is a test statistic of the Bagdonavičus-Nikulin test based on the MCE model (the test statistic represented in [6]), S_{BN3} is a test statistic of the Bagdonavičus-Nikulin test based on the SCE model (the test statistic represented in [7]), $F_{N(0,1)}(t)$ is a CDF of the standard normal distribution at time t and $F_{\chi^2(k)}(t)$ is a CDF of the chi-square distribution with k degrees of freedom at time t.

The $MIN3$ test has a left-side critical area.

The asymptotical distribution of the $MIN3$ test statistic is unknown. However, its distribution can be approximated by the Beta distribution of the third kind distribution that is represented in next section.

3 Distribution of the $MIN3$ Test Statistic Under Null Hypothesis

Because the p-value of any test is distributed asymptotically uniformly on the interval $[0, 1]$ under the null hypothesis H_0, we can suppose that the distribution $G_n(S_{MIN3}|H_0)$ of the test statistic S_{MIN3} that is a minimum of three dependent random variables (p-values of the WKM, BN2 and BN3 tests) converges to the limit distribution $G(S_{MIN3}|H_0)$ asymptotically.

3.1 Simulation of the Distribution of the $MIN3$ Test Statistic

The distributions of the $MIN3$ test statistic were simulated by the Monte Carlo method with $N = 2\ 700\ 000$ replications. It makes possible to conclude that the difference [8] $D_N = \sup_t \left|\hat{F}_N(t) - F_{S_{MIN3}}(t)\right|$ between empirical CDF $\hat{F}_N(t)$ and CDF of the $MIN3$ test statistic is not greater than 0.001 with a confidence probability 0.99. In Figures 1 and 2, the distributions of the $MIN3$ test statistic $G(S_n|H_0)$ are shown in the uncensored and censored cases. On the graphics, we can see that the CDF of the $MIN3$ test statistic tends to the limit distribution with growing n, however this limit distribution is unknown now. If the sample includes the censored observations, then the distribution of the $MIN3$ test statistic is slightly changed.

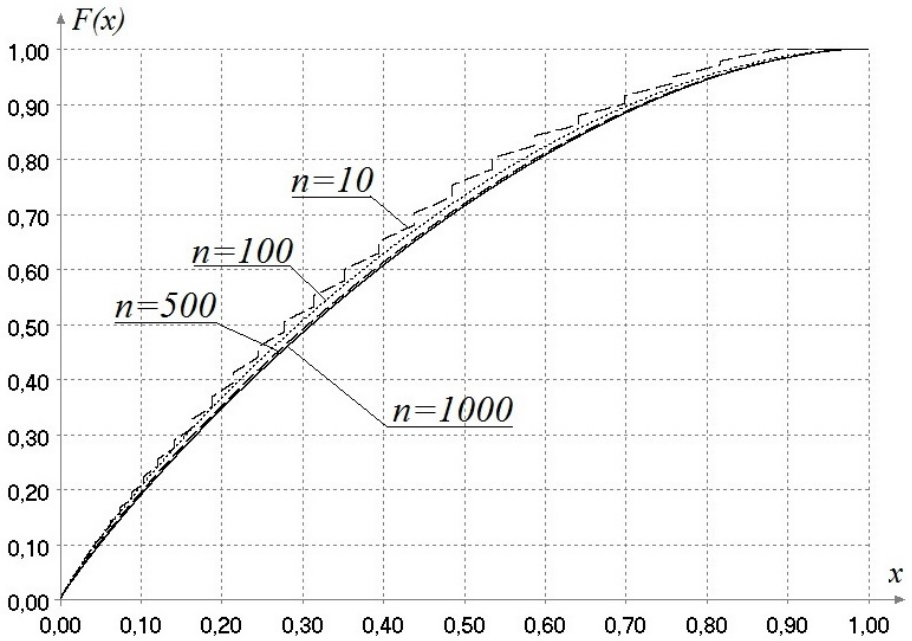


Figure 1: CDF $G_n(S|H_0)$ of the $MIN3$ test statistic S_{MIN3} for various samples size $n_1 = n_2 = n$ in the uncensored case

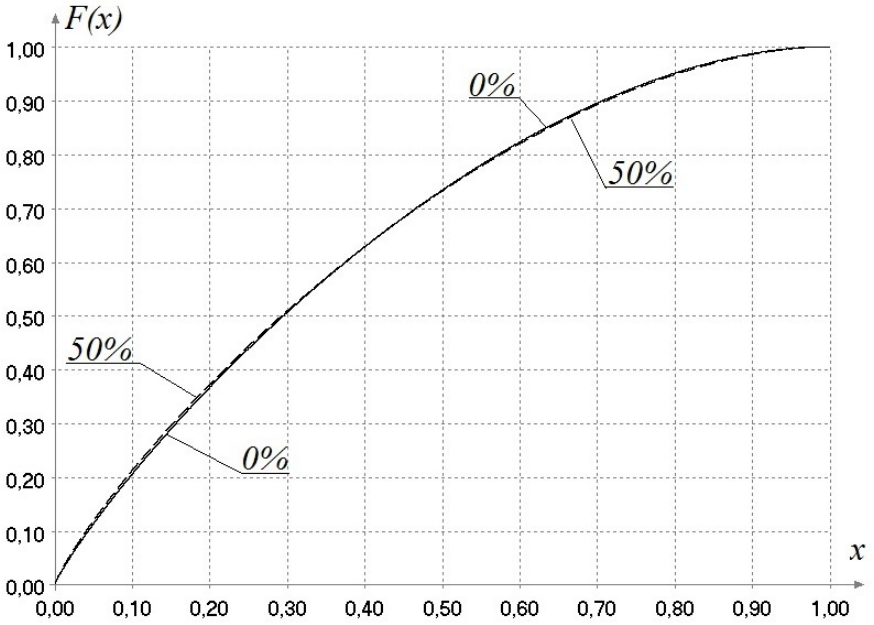


Figure 2: CDF $G_n(S|H_0)$ of the $MIN3$ test statistic S_{MIN3} for sample size $n_1 = n_2 = 100$ in the censored case

3.2 Approximation of the Distribution of the $MIN3$ Test Statistic

The approximation of the test statistic distribution is a possible way to the more correct application of the statistical test [9]. To approximate the distribution of the $MIN3$ test statistic, we are looking for a distribution from the class of the Beta-Generated distributions that can be defined as

$$F(x;\alpha,\beta) = \frac{B(G(x),\alpha,\beta)}{B(\alpha,\beta)},$$

where $B(\alpha,\beta) = \int\limits_0^1 y^{\alpha-1}(1-y)^{\beta-1}dy$ and $B(x,\alpha,\beta) = \int\limits_0^x y^{\alpha-1}(1-y)^{\beta-1}dy$ are complete and incomplete beta functions, α and β are parameters, and $G(x)$ is a generating CDF of some random variable that can also have own parameters. Using the different generating functions $G(x)$ we can obtain a lot of flexible distributions, for example,

- Beta distribution of the first kind, $G(x) = x,\ 0 \le x \le 1$;
- Beta distribution of the second kind, $G(x) = \frac{x}{1+x},\ 0 \le x < +\infty$;
- Beta distribution of the third kind [10], $G(x,\delta) = \frac{\delta x}{1+(\delta-1)x},\ 0 \le x < 1$;
- Generalized Beta distribution [11], $G(x,\delta,\gamma) = \frac{x^\gamma}{1+\delta x^\gamma},\ 0 \le x^\gamma \le \frac{1}{(1-\delta)}$;
- Pareto's distribution, $G(x) = \frac{x-1}{x},\ 1 \le x < \infty,\ \alpha = 1$;

- Exponential generalized beta distribution of the second type, $G(x) = \frac{e^x}{1+e^x}$, $-\infty < x < +\infty$.

Since the $MIN3$ test statistic is defined in the interval $[0, 1]$ the Beta distribution of the first and the third kind are the most suitable. However, the Beta distribution of the first kind has only two parameters, while the Beta distribution of the third kind has three parameters and is better suited for the approximation. The distribution of the $MIN3$ test statistic can be good approximated by the Beta distribution of the third kind with a probability density function (PDF) [10]

$$f_{\text{Beta}_{\text{III}}}(x; a, b, c) = \frac{c^a}{B(a, b)} x^{a-1}(1-x)^{b-1}(1+(c-1)x)^{-a-b}, x \in (0, 1), a, b, c > 0.$$

The estimated parameters $\hat{a}$,$\hat{b}$ and $\hat{c}$ of the Beta distribution of the third kind and Kolmogorov's distance [8] $D_{n,N} = \sup_t \left|\hat{F}_{n,N}(t) - F_{BetaIII}(t; a, b, c)\right|$ between the empirical CDF $\hat{F}_{n,N}(t)$ of the $MIN3$ test statistic and their approximations $F_{BetaIII}(t; a, b, c)$ for different n are shown in Table 1. The estimations were found by the software system "ISW" [12]. As long as the sample size $n_1 = n_2 = n$ increases, the distance $D_{n,N}$ is reduced and stabilized close to the simulation error 0.001. Using the approximation, we can draw a PDF of the $MIN3$ test statistic that is shown in Figure 3.

Table 1: Approximation of the $MIN3$ test statistic distribution using the Beta distribution of the third kind

n	$\hat{a}$	$\hat{b}$	$\hat{c}$	$D_{n,N}$
10	0.8603	2.8066	0.5579	0.0204
20	0.8308	2.2859	0.6813	0.0071
50	0.8576	2.0410	0.7957	0.0023
100	0.8769	1.9175	0.8583	0.0017
200	0.8859	1.8397	0.8908	0.0008
500	0.9020	1.7946	0.9172	0.0007
1000	0.9086	1.8007	0.9031	0.0008

3.3 Lower Percentage Points of the $MIN3$ Test Statistic Distribution

In Table 2, the lower percentage points of the distribution $G_n(S_{MIN3}|H_0)$ for samples size $n = n_1 = n_2$ in the range from 10 to 1000 observations without censoring are shown. In Table 3, the distributions of the $MIN3$ test statistic $G_n(S_{MIN3}|H_0)$ for samples size $n_1 = n_2 = 100$ and censoring rates in the range from 0% to 50%. Thus, if a researcher has two samples with corresponding sizes $n = n_1 = n_2$, then the lower percentage points in Tables 2 and 3 can be used for rejecting of the null hypothesis.

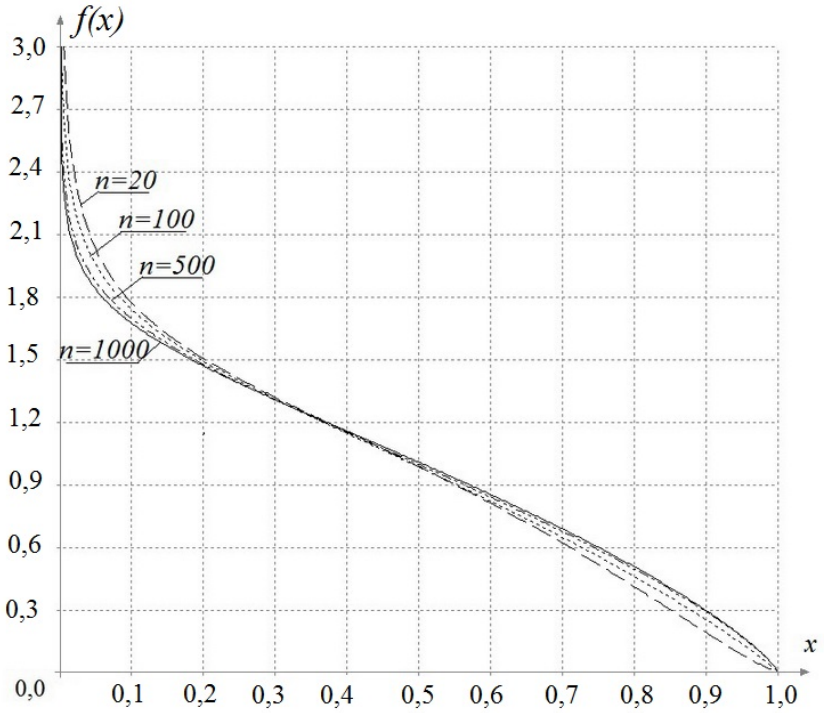


Figure 3: PDF $g_n(S|H_0)$ of the $MIN3$ test statistic S_{MIN3} for sample size $n_1 = n_2 = n$ without censoring

In the general case, the p-value can be calculated using the simulation procedure by the Monte Carlo method.

Conclusion

Using the Monte-Carlo simulation, it is possible to estimate the limit distribution of the $MIN3$ test statistic. Special software ISW allows us to establish the closest family of distributions describing the distribution of the $MIN3$ test statistic, as well as to estimate the parameters of this family. In the paper, the parameters of the distribution of $MIN3$ test statistic are represented both for a non-censoring case ($r = r(X_1) = r(X_2), r = 0\%$) and for a censoring case ($r = r(X_1) = r(X_2), 10\% \leq r \leq 50\%$).

The Kolmogorov's distance between the empirical CDF of the $MIN3$ test statistic and their approximations $F_{BetaIII}(t; a, b, c)$ for different n are shown in the paper.

The analytic form of the asymptotic distribution of the $MIN3$ test statistic is a next step of this test research. Nonetheless the approximation of the distribution of the $MIN3$ test statistic and the tables of the percentage points can be used right now.

The results obtained in the paper can be used as recommendations for using of the $MIN3$ test in practice.

Table 2: Lower percentage points of the $MIN3$ test statistic distribution under H_0 without censoring

α	$G^{-1}_{S_{MIN_3}\|H_0}(\alpha)$						
	$n = 10$	$n = 20$	$n = 50$	$n = 100$	$n = 200$	$n = 500$	$n = 1000$
0.001	0.0003	0.0002	0.0002	0.0002	0.0002	0.0003	0.0003
0.005	0.0017	0.0013	0.0013	0.0015	0.0016	0.0018	0.0019
0.010	0.0038	0.0029	0.0029	0.0032	0.0034	0.0038	0.0040
0.020	0.0077	0.0066	0.0066	0.0070	0.0074	0.0080	0.0085
0.030	0.0119	0.0104	0.0105	0.0111	0.0117	0.0126	0.0131
0.040	0.0162	0.0146	0.0146	0.0154	0.0162	0.0172	0.0179
0.050	0.0200	0.0185	0.0188	0.0198	0.0208	0.0220	0.0229
0.060	0.0245	0.0230	0.0232	0.0244	0.0255	0.0269	0.0280
0.070	0.0299	0.0273	0.0278	0.0290	0.0303	0.0319	0.0331
0.080	0.0344	0.0321	0.0323	0.0337	0.0351	0.0369	0.0383
0.090	0.0379	0.0367	0.0371	0.0386	0.0402	0.0420	0.0436
0.100	0.0438	0.0412	0.0419	0.0435	0.0454	0.0473	0.0489

References

[1] S. Postovalov, P. Philonenko (2013). A Comparison of Homogeneity Tests for Different Alternative Hypotheses, Statistical Models and Methods for Reliability and Survival Analysis : monograph. - London : Wiley-ISTE. - Chap. 12. - P. 177–194. - (Mathematics and Statistics series).

[2] Petr Philonenko, Sergey Postovalov (2017). The Convergence Rate Research of Two-Sample Test Statistic Distribution to the Limit Distribution under Right-Censored Data / Obrabotka informacii i matematicheskoe modelirovanie [Information processing and mathematical simulation], (April, 25–26, 2017, Novisibirsk). - Novosibirsk : SibGUTI, 2017. - pp. 187–193. - ISBN 978-5-31434-038-1. (In Russian).

[3] Petr Philonenko and Sergey Postovalov (2019). The new robust two-sample test for randomly right-censored data, Journal of Statistical Computation and Simulation, DOI: 10.1080/00949655.2019.1578769.

[4] Metropolis, N.; Ulam, S. (1949). "The Monte Carlo Method". Journal of the American Statistical Association. 44 (247): 335–341. doi:10.2307/2280232. JSTOR 2280232. PMID 18139350.

[5] Pepe M.S. and Fleming T.R., Weighted Kaplan-Meier statistics: a class of distance tests for censored survival data. Biometrics. 1989; 45 : 497–507.

[6] Bagdonavičius V.B., Nikulin M., "On goodness-of-fit tests for homogeneity and proportional hazards", Applied Stochastic Models in Business and Industry, vol. 22, no. 1, pp. 607–619, 2006.

Table 3: Lower percentage points of the $MIN3$ test statistic distribution under H_0 with the censoring rate up to 50%, $n = n_1 = n_2 = 100$

α	$G^{-1}_{S_{MIN_3}\|H_0}(\alpha)$					
	0%	10%	20%	30%	40%	50%
0.001	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
0.010	0.0032	0.0031	0.0030	0.0030	0.0029	0.0028
0.020	0.0070	0.0068	0.0067	0.0066	0.0064	0.0063
0.030	0.0111	0.0108	0.0106	0.0104	0.0102	0.0100
0.040	0.0153	0.0150	0.0147	0.0145	0.0142	0.0140
0.050	0.0197	0.0193	0.0190	0.0189	0.0183	0.0182
0.060	0.0243	0.0238	0.0235	0.0232	0.0227	0.0225
0.070	0.0289	0.0285	0.0280	0.0278	0.0272	0.0269
0.080	0.0337	0.0332	0.0327	0.0324	0.0319	0.0314
0.090	0.0385	0.0380	0.0374	0.0371	0.0366	0.0360
0.100	0.0435	0.0428	0.0424	0.0420	0.0414	0.0408

[7] Bagdonavičius, V., Kruopis, J. and Nikulin, M. S. (2013) Censored and Truncated Data, in Non-parametric Tests for Censored Data, John Wiley & Sons, Inc, Hoboken, NJ, USA.

[8] Kolmogorov A (1933). "Sulla determinazione empirica di una legge di distribuzione". G. Ist. Ital. Attuari. 4: 83–91.

[9] Lemeshko B. Yu., Lemeshko S. B., Postovalov S. N. (2010). Statistic Distribution Models for Some Nonparametric Goodness-of-Fit Tests in Testing Composite Hypotheses, Communications in Statistics – Theory and Methods, 39:3, 460–471.

[10] Gubarev V.V. Probabilistic models: Reference book in 2 parts. // NETI. - Novosibirsk, 1992. - 422 P. [in Russian].

[11] McDonald, James B., Xu, Yexiao J. (1995) "A generalization of the beta distribution with applications". Journal of Econometric, 66(1-2), 133–152. doi:10.1016/0304-4076(94)01612-4.

[12] Chimitova E.V., Lemeshko B.Yu., Lemeshko S.B., Postovalov S.N., Rogozhnikov A.P., "Software system for simulation and research of probabilistic regularities and statistical data analysis in reliability and quality control", Mathematical and Statistical Models and Methods in Reliability: Applications to Medicine, Finance, and Quality Control, Statistics for Industry and Technology, 2010, 417–432.

The research of the two-sample test statistics convergence rate

Petr Philonenko and Sergey Postovalov
Novosibirsk State Technical University, Novosibirsk, Russia
e-mail: petr-filonenko@mail.ru, postovalov@ngs.ru

Abstract

In the paper, we present the results of the computer simulation for the convergence rate of two-sample test statistic distributions in the lifetime data case. Various distributions of failure times and various distributions of censored times are considered. In the result the dependence on the Kolmogorov's distance between a distribution of the two-sample test statistic and its limit distribution is shown.

Keywords: Monte-Carlo method, survival analysis, hypothesis testing, randomly right-censored observations, convergence rate, Gehan's Generalized Wilcoxon Test, Peto and Peto's Generalized Wilcoxon Test, log-rank test, Cox-Mantel test, Bagdonavičius-Nikulin tests, weighted log-rank tests, Kolmogorov's distance.

Introduction

When conducting a procedure for hypothesis testing, for example, the hypothesis of homogeneity [1], it is necessary to compare the computed value of the test statistic with a certain critical value. Such critical value represents a certain quantile of the distribution of the test statistic. However under small sample size, the distribution of test statistic can differ significantly from the limit distribution [2] which can lead to wrong decision. Therefore, the goal of our work is to study distributions of test statistic for various sample sizes and define the difference (in the sense of the Kolmogorov's metric [3]) to the corresponding limit distribution using the Monte-Carlo simulation.

In Section 2, we present the statement of the two-sample problem and considered two-sample tests under randomly right-censored observations. Also in Section 2, we give a methodology for studying the rate of convergence of the test statistic distribution to the corresponding limit distribution. In Section 3, we present the results of computer simulation of the convergence rate study.

1 Problem Statement

Suppose that we have two samples of continues variables ξ_1 and ξ_2 respectively, $X_1 = \{t_{11}, t_{12}, ..., t_{1n_1}\}$ and $X_2 = \{t_{21}, t_{22}, ..., t_{2n_2}\}$ of two survival distributions $S_1(t)$

This research has been supported by the Russian Ministry of Education and Science as a part of the state task (application number 1.1009.2017/4.6).

and $S_2(t)$. The samples size are n_1, n_2 (if $n_1 = n_2$, then may be denoted as n). The observation $t_{ij} = \min(T_{ij}, C_{ij})$, where T_{ij} and C_{ij} are the failure and censoring times for the j-th object of the i-th group. T_{ij} and C_{ij} are i.i.d. with a cumulative distribution function (CDF) $F_i(t)$ and $F_i^C(t)$ respectively. Survival curve means the probability of survival in the time interval $(0, t)$

$$S_i(t) = P\{\xi_i > t\} = 1 - F_i(t).$$

Then the null hypothesis is

$$H_0 : S_1(t) = S_2(t)$$

against alternative hypothesis

$$H_1 : S_1(t) \neq S_2(t).$$

Let an indicator of censoring c_{ij} be equal 0 if t_{ij} is a failure time and be equal 1 if t_{ij} is a censored time.

Censoring rate r for sample X_i is

$$r(X_i) = \left(\sum_{j=1}^{n_i} c_{ij}\right) \times \frac{100\%}{n_i}, i = 1, 2.$$

2 Convergence Rate of Test Statistic Distribution to the Limit Distribution

Let a test statistic S_n has (when the hypothesis H_0 is true) a pre-limit distribution $G_n(t)$ when sample sizes are $n < \infty$ and S_n has the limit distribution $G(t)$ when $n \to \infty$. It is impossible to prove a convergence of $G_n(t)$ to $G(t)$ using simulation but one can determine a sample size n providing a maximal distance between the limit and pre-limit distributions no more than ϵ, for example, $\epsilon = 0.01$.

Finding an analytic form of the pre-limit distribution $G_n(t)$ is usually more difficult task than finding the limit distribution $G(t)$. However, one can estimate empirical distribution function of $G_n(t)$ using the Monte Carlo method. It is necessary simulate a sample (with a sample size N) contained test statistic values S_n: $\left\{S_n^{(1)}, S_n^{(2)}, ..., S_n^{(N)}\right\}$ and compute empirical distribution function $G_{n,N}(t)$. Hence, the Kolmogorov's distance [4],[5] is computed as

$$D_{n,N} = \sup_{|t|<\infty} |G_{n,N}(t) - G(t)|.$$

Surely, $G_{n,N}(t)$ differs from $G_n(t)$ with accurate to ϵ_N. But using the Kolmogorov's theorem [6], one can determine N that $\epsilon_N < \epsilon$.

According to the Kolmogorov's theorem

$$\lim_{N\to\infty} P\left\{\sqrt{N} \sup_{|x|<\infty} |G_{n,N}(x) - G_n(x)| < t\right\} = K(t),$$

where $K(t)$ is the Kolmogorov's distribution. One can find such N providing $\epsilon_N < \epsilon$ with a some confidence probability. For instance, if we want a simulation error $\epsilon_N \leq 0.001$ with the confidence probability 0.99, then required sample size N is

$$N = \left[\left(\frac{K^{-1}(0.99)}{0.001}\right)^2\right] + 1 \approx \left[\left(\frac{1.62762}{0.001}\right)^2\right] + 1 = 2\ 649\ 147.$$

3 Two-Sample Tests

In the paper, we consider following statistical tests for the solution of the two-sample problem:

1. the Gehan's Generalized Wilcoxon Test ([7], [8]);
2. the Peto and Peto's Generalized Wilcoxon Test ([8], [9]);
3. the log-rank test ([8], [10]);
4. the Cox-Mantel test ([8], [11]);
5. the Bagdonavičius-Nikulin tests based on the generalized Cox model [12];
6. the Bagdonavičius-Nikulin test based on the SCE model [13];
7. the Tarone-Ware test (weighted log-rank test) [13].

Additional information of these test statistics can be founded in presented papers.

4 Simulation

In the study of the distance between pre-limit $G_{n,N}(t)$ and limit $G(t)$ test statistic distributions, we consider $\varepsilon = 0.01$, $N = 2.7 \cdot 10^6$ replications ($\varepsilon_N \leq 0.001$), the laws of failure time distributions are Weibull-Gnedenko and Exponential, the laws of censoring time distribution are Weibull-Gnedenko and Gamma, censoring rate is in the range from 0% and 50% (randomly right-censored observations). The results of computer simulation are given in Tables 1–7.

From the obtained results, it is obvious that in the case of the Gehan test, if the sample size $n \geq 20$ observations, the Kolmogorov's distance does not exceed 0.01 both in the case of complete observations and in the case with a censoring rate 50%. For the Peto, log-rank and Cox-Mantel two-sample tests, in the case of complete data, if the sample size $n \geq 20$ observations, the Kolmogorov's distance to the limit distributions does not exceed 0.01, and in the case with censoring rate 50% a similar result is achieved at $n \geq 10$.

For the Tarone-Ware two-sample test, if the sample size $n \geq 20$ observations, the Kolmogorov's distance to the limit distribution is no more than 0.01.

Table 1: Simulated Kolmogorov's distance between pre-limit and limit distributions of the Gehan's Generalized Wilcoxon test statistic

$n = n_1 = n_2$	$F \sim We$, $F^C \sim We$		$F \sim We$, $F^C \sim \Gamma$		$F \sim Exp$, $F^C \sim We$		$F \sim Exp$, $F^C \sim \Gamma$	
	0%	**50%**	**0%**	**50%**	**0%**	**50%**	**0%**	**50%**
10	0.0171	0.0227	0.0172	0.0114	0.0173	0.0328	0.0171	0.0126
20	0.0071	0.0075	0.0070	0.0039	0.0069	0.0113	0.0068	0.0040
30	0.0041	0.0045	0.0041	0.0022	0.0040	0.0061	0.0041	0.0019
40	0.0030	0.0030	0.0029	0.0014	0.0028	0.0044	0.0026	0.0016
50	0.0024	0.0018	0.0023	0.0011	0.0022	0.0030	0.0024	0.0015

Table 2: Simulated Kolmogorov's distance between pre-limit and limit distributions of the Peto and Peto's Generalized Wilcoxon test statistic

$n_1 = n_2$	$F \sim We$, $F^C \sim We$		$F \sim We$, $F^C \sim \Gamma$		$F \sim Exp$, $F^C \sim We$		$F \sim Exp$, $F^C \sim \Gamma$	
	0%	**50%**	**0%**	**50%**	**0%**	**50%**	**0%**	**50%**
10	0.0176	0.0058	0.0172	0.0050	0.0172	0.0063	0.0169	0.0048
20	0.0065	0.0030	0.0068	0.0025	0.0069	0.0028	0.0069	0.0024
30	0.0039	0.0018	0.0039	0.0016	0.0042	0.0018	0.0042	0.0022
40	0.0029	0.0020	0.0026	0.0014	0.0028	0.0019	0.0030	0.0015
50	0.0023	0.0013	0.0025	0.0009	0.0022	0.0014	0.0023	0.0012

Table 3: Simulated Kolmogorov's distance between pre-limit and limit distributions of the log-rank test statistic

$n_1 = n_2$	$F \sim We$, $F^C \sim We$		$F \sim We$, $F^C \sim \Gamma$		$F \sim Exp$, $F^C \sim We$		$F \sim Exp$, $F^C \sim \Gamma$	
	0%	**50%**	**0%**	**50%**	**0%**	**50%**	**0%**	**50%**
10	0.0144	0.0071	0.0136	0.0046	0.0138	0.0084	0.0139	0.0056
20	0.0075	0.0038	0.0076	0.0025	0.0075	0.0045	0.0075	0.0027
30	0.0056	0.0028	0.0055	0.0022	0.0055	0.0033	0.0050	0.0018
40	0.0041	0.0020	0.0041	0.0015	0.0042	0.0025	0.0045	0.0019
50	0.0035	0.0020	0.0034	0.0014	0.0036	0.0024	0.0035	0.0013

Table 4: Simulated Kolmogorov's distance between pre-limit and limit distributions of the Cox-Mantel test statistic

$n_1 = n_2$	$F \sim We$, $F^C \sim We$		$F \sim We$, $F^C \sim \Gamma$		$F \sim Exp$, $F^C \sim We$		$F \sim Exp$, $F^C \sim \Gamma$	
	0%	**50%**	**0%**	**50%**	**0%**	**50%**	**0%**	**50%**
10	0.0127	0.0068	0.0123	0.0044	0.0125	0.0087	0.0127	0.0051
20	0.0072	0.0037	0.0069	0.0021	0.0071	0.0043	0.0068	0.0028
30	0.0050	0.0024	0.0047	0.0018	0.0049	0.0031	0.0049	0.0018
40	0.0042	0.0021	0.0036	0.0013	0.0038	0.0025	0.0036	0.0017
50	0.0032	0.0019	0.0031	0.0012	0.0032	0.0022	0.0034	0.0014

Table 5: Simulated Kolmogorov's distance between pre-limit and limit distributions the Tarone-Ware test statistic (weighted log-rank test)

$n_1 = n_2$	$F \sim We$, $F^C \sim We$		$F \sim We$, $F^C \sim \Gamma$		$F \sim Exp$, $F^C \sim We$		$F \sim Exp$, $F^C \sim \Gamma$	
	0%	**50%**	**0%**	**50%**	**0%**	**50%**	**0%**	**50%**
10	0.0108	0.0107	0.0109	0.0085	0.0104	0.0121	0.0105	0.0093
20	0.0056	0.0048	0.0053	0.0043	0.0054	0.0054	0.0056	0.0046
30	0.0034	0.0028	0.0032	0.0028	0.0035	0.0037	0.0038	0.0028
40	0.0031	0.0029	0.0028	0.0015	0.0025	0.0027	0.0024	0.0022
50	0.0020	0.0024	0.0026	0.0015	0.0024	0.0020	0.0019	0.0014

Table 6: Simulated Kolmogorov's distance between pre-limit and limit distributions of test statistic of the Bagdonavičius-Nikulin test based on the SCE model

$n_1 = n_2$	$F \sim We$, $F^C \sim We$		$F \sim We$, $F^C \sim \Gamma$		$F \sim Exp$, $F^C \sim We$		$F \sim Exp$, $F^C \sim \Gamma$	
	0%	**50%**	**0%**	**50%**	**0%**	**50%**	**0%**	**50%**
10	0.0657	0.0472	0.0657	0.0304	0.0656	0.0577	0.0653	0.0357
20	0.0437	0.0264	0.0438	0.0132	0.0434	0.0357	0.0436	0.0178
30	0.0337	0.0196	0.0337	0.0094	0.0333	0.0275	0.0334	0.0123
40	0.0275	0.0158	0.0273	0.0063	0.0282	0.0226	0.0274	0.0098
50	0.0240	0.0136	0.0240	0.0053	0.0238	0.0202	0.0237	0.0082

Table 7: Simulated Kolmogorov's distance between pre-limit and limit distributions of test statistic of the Bagdonavičius-Nikulin test based on the generalized Cox model

$n_1 = n_2$	$F \sim We$, $F^C \sim We$		$F \sim We$, $F^C \sim \Gamma$		$F \sim Exp$, $F^C \sim We$		$F \sim Exp$, $F^C \sim \Gamma$	
	0%	**50%**	**0%**	**50%**	**0%**	**50%**	**0%**	**50%**
10	0.0578	0.0455	0.0585	0.0294	0.0584	0.0555	0.0574	0.0348
20	0.0331	0.0245	0.0334	0.0129	0.0334	0.0317	0.0328	0.0173
30	0.0247	0.0173	0.0239	0.0086	0.0242	0.0234	0.0240	0.0113
40	0.0193	0.0138	0.0191	0.0060	0.0184	0.0194	0.0193	0.0089
50	0.0159	0.0114	0.0161	0.0050	0.0155	0.0164	0.0160	0.0076

For the Bagdonavičius-Nikulin tests, the results of the convergence rate are not so ambiguous and may depend on the censored time distribution. However, if the sample size $n \geq 200$ (for two-sample test based on the SCE model), and $n \geq 100$ (for a two-sample test based on generalized Cox model), the Kolmogorov's distance to the corresponding limit distribution does not exceed 0.01.

The results obtained can be used as recommendations on the application of these criteria in practice.

Conclusions

Summing up, we note that in this paper we described a procedure on the basis of which we can define such sample size that the distance between prelimit and limit distributions does not exceed a certain value. It means the application of the limit distribution for hypothesis testing is possible without loosing the accuracy in the statistical conclusions. For this procedure, we conducted a computer simulation using two-sample tests for both complete and censored observations. The obtained results can be used as recommendations, starting from such volume the distance of the pre-limit distribution to the limit distribution does not exceed 0.01. If the sample size is not enough, then we can recommend to simulate the test statistics distribution using the Monte-Carlo method.

The work was done with the financial support of the Ministry of Education and Science of the Russian Federation as part of the project part of the state task (application number 1.1009.2017/4.6).

References

[1] Petr Philonenko, Sergey Postovalov (2015) A new two-sample test for choosing between log-rank and Wilcoxon tests with right-censored data, Journal of Statistical Computation and Simulation, 85:14, 2761-2770.

[2] Lemeshko S.B., Nikulin M.S., Saaidia N. Simulation of the statistics distributions and power of the goodness-of-fit tests in composite hypotheses testing rather Inverse Gaussian distribution // Proc. 6th St. Petersburg Workshop on Simulation. St. Petersburg. 2009. Vol. 1, 323–328.

[3] Chimitova E. Application of classical Kolmogorov, Cramer-von Mises-Smirnov and Anderson-Darling tests for censored samples / E. Chimitova, H. Liero, M. Vedernikova // Proceedings of the International Workshop "Applied Methods of Statistical Analysis. Simulations and Statistical Inference". - Novosibirsk: Publ. house of NSTU, 2011. - pp. 176–185.

[4] Ishalina M.A. Convergence of two-sample test statistic distributions to the limiting law / M.A. Ishalina, S.N. Postovalov // Proceedings Third International Conference on Accelerated Life Testing, Reliability-based Analysis and Design. 19-21 May 2010, Clermont-Ferrand, France. - pp. 237–242.

[5] Smirnov N. V. "Tables for estimating the goodness of fit of empirical distributions", Annals of Mathematical Statistics, vol. 19, num. 1, 1948, pp. 279.

[6] Kobzar A. I. Applied mathematical statistcs. For engineers and scientifical reseracher. - Moscow. Fizmatlit. 2006. - P. 816. (In Russian).

[7] Gehan E.A., A generalized Wilcoxon test for comparing arbitrarily singly-censored sam-ples // Biometrika. 1965. V. 52, N 1/2. P. 203-223.

[8] Lee E.T., Wang J.W., Statistical methods for survival data analysis. 3rd ed. Hoboken (NJ): John Wiley & Sons; 2003. doi:10.1002/0471458546.index.

[9] Peto R., Peto J., Asymptotically efficient rank invariant test procedures//J. Royal Statist. Soc. Ser. A (General). 1972. V. 135, N 2. P. 185-207.

[10] Savage, I. R. (1956). Contributions to the Theory of Rank Order Statistics: The Two Sample Case. Annals of Mathematical Statistics, 27, 590-615.

[11] Mantel, N. (1966). Evaluation of Survival Data and Two New Rank Order Statistics Arising in Its Consideration. Cancer Chemotherapy Reports, 50, 163-170.

[12] Bagdonavičius V. B., Levuliene R. J., Nikulin M. S., Zdorova-Cheminadeo, "Tests for equality of survival distributions against nonlocation alternatives", Lifetime Data Analysis, vol. 10, num. 4, 2004, p. 445-460.

[13] Bagdonavičius, V., Kruopis, J. and Nikulin, M. S. (2013) Censored and Truncated Data, in Non-parametric Tests for Censored Data, John Wiley & Sons, Inc, Hoboken, NJ, USA.

Optimal index estimation of log-gamma distribution

Dimitris N. Politis[1], Vyacheslav A. Vasiliev[2] and Sergey E. Vorobeychikov[2]
University of California, San Diego, USA[1],
National Research Tomsk State University, Tomsk, Russia[2]
e-mail: dpolitis@ucsd.edu, vas@mail.tsu.ru, sev@mail.tsu.ru

Abstract

The problem of optimal estimation of the heavy tail index is revisited from the point of view of truncated estimation. A class of these estimators is introduced having guaranteed accuracy based on a sample of fixed size [7]. The optimality of considered log-gamma index estimators in the sense of a special type risk function is established. The considered risk function makes possible to optimize not only the asymptotic variances of the estimators, as well as used for estimation of sample size. Optimization of the parameters of log-gamma distribution is presented. Simulation results confirm theoretical one's.

Keywords: Optimal parameter estimation, heavy tails, log-gamma distribution, optimal convergence rate.

Introduction

This paper presents results of optimality for the parameter estimators of log-gamma distribution, introduced in [7]. Some general properties of parameter estimators are used only and are such that the considered class of estimators is sufficiently wide.

In this paper, we use the risk function of a special type which is a linear combination of mean-square deviation of parameter estimators and sample size. The requirement of both good parameter estimation quality and reasonable duration of observations is formulated as a risk efficiency problem. The risk function of similar structure was proposed in [1], see also references therein. The criterion is given by a certain loss function and optimization is performed based on it.

Further the loss and risk functions of the type proposed in [1] were used in, e.g., [8, 9] for optimization of interpolators and predictors of a scalar AR(1) process with unknown parameters. Similar optimization problem of the sequential parameter estimator of AR(1) was considered in [3]. There was considered a risk function defined on the basis of squared estimation error of sequential estimator of the dynamic parameter.

Later the results of those papers were refined and extended to other stochastic models. In particular, this approach was applied to construction of optimal adaptive predictors of the stochastic processes related with discrete and continuous-time dynamical systems, see, e.g, [16, 2]. The proposed procedures are based on the so-called truncated estimators which have been developed in order to estimate ratio type functionals from a wide class by dependent observations and by samples of fixed size so that they had guaranteed accuracy in the sense of the L_{2m}-norm, $m \geq 1$. Examples of parameter estimation problems of discrete and continuous time systems on a time interval of a fixed length are considered.

The truncated estimators may keep asymptotic properties of the estimators they are based upon. Another approaches do not guarantee prescribed estimation accuracy when using samples of non-random finite size and lead up to complicated analytical problems in adaptive procedures. Applications of truncated estimators with the said quality makes possible to optimize the risk function which is a linear combination of sample mean of mean-square deviation of predictors and sample size.

Results of non-asymptotic non-parametric problems can be found also in [5, 6] among others. In particular, they have investigated non-asymptotic properties of the regression and density function kernel-type estimators.

It should be noted that first truncated parameter estimation method was applied for construction of adaptive optimal predictors of VAR(1) in [10]. Then this method was applied to more complicated stochastic systems. Among the processes considered are stable multivariate discrete time AR(1), ARMA(1,1) and RCA(1), as well as continuous time diffusion and time delayed processes, see, e.g., [2]. The proposed procedure is shown to be asymptotically risk efficient as the cost of prediction error tends to infinity.

1 Log-gamma density function

Consider the parameter estimation problem based on i.i.d. observations $X_1, \ldots, X_n$ with the log-gamma density function

$$f(x) = C_f x^{-(\gamma+1)} \log^{\beta-1} x, \quad x \geq 1.$$

Our main aim is to prove the optimality of the truncated estimators β_n, γ_n and θ_n of the parameters β, γ and θ presented in [7] in the sense of the risk function considered above.

To define the truncated estimators we introduce, similar to [7] for some given $a > 0$ the functional

$$\Phi(a) = E \log^a X_1.$$

Using the definition of $f(x)$ according to [7] we have

$$\Phi(a) = \frac{\gamma}{\beta + a} \Phi(a + 1).$$

Analogously for a given $b \neq a$,

$$\Phi(b) = \frac{\gamma}{\beta + b} \Phi(b + 1).$$

Thus

$$\beta\Phi(a) - \gamma\Phi(a + 1) = -a\Phi(a),$$
$$\beta\Phi(b) - \gamma\Phi(b + 1) = -b\Phi(b)$$

and the solution of this system has the form

$$\beta = \frac{b\Phi(b)\Phi(a + 1) - a\Phi(a)\Phi(b + 1)}{\Delta_{a,b}},$$

$$\gamma = \frac{(b-a)\Phi(a)\Phi(b)}{\Delta_{a,b}},$$

as well as

$$\theta = \frac{\Delta_{a,b}}{(b-a)\Phi(a)\Phi(b)},$$

where

$$\Delta_{a,b} = \Phi(a)\Phi(b+1) - \Phi(b)\Phi(a+1).$$

Now we define the empirical functional estimator

$$\Phi_n(a) = \frac{1}{n}\sum_{k=1}^{n} \log^a X_k$$

of $\Phi(a)$ and the truncated estimators β_n, γ_n and θ_n (see also [7]) as follows

$$\beta_n = \frac{b\Phi_n(b)\Phi_n(a+1) - a\Phi_n(a)\Phi_n(b+1)}{\Delta_{a,b}(n)} \cdot \chi(|\Delta_{a,b}(n)| \geq \log^{-1} n), \tag{1}$$

$$\gamma_n = \frac{(b-a)\Phi_n(a)\Phi_n(b)}{\Delta_{a,b}(n)} \cdot \chi(|\Delta_{a,b}(n)| \geq \log^{-1} n), \tag{2}$$

$$\theta_n = \frac{\Delta_{a,b}(n)}{(b-a)\Phi_n(a)\Phi_n(b)} \cdot \chi(|(b-a)\Phi_n(a)\Phi_n(b)| \geq \log^{-1} n), \tag{3}$$

where

$$\Delta_{a,b}(n) = \Phi_n(a)\Phi_n(b+1) - \Phi_n(b)\Phi_n(a+1).$$

From [7] it follows that the asymptotic normality property, defined in [7] is fulfilled for the estimators γ_n, θ_n and β_n with the rate $\alpha_n = \sqrt{n}$ and the asymptotic variance of $n \cdot \gamma_n$ is defined by equations

$$\sigma_\gamma^2 = (b-a)^2\Delta_{a,b}^{-2} \cdot \sigma_1^2 + 2(b-a)^2\Delta_{a,b}^{-3} \cdot \sigma_2 + (b-a)^2\Delta_{a,b}^{-4} \cdot \sigma_3^2, \tag{4}$$

where

$$\sigma_1^2 = \Phi^2(a)\Phi(2b) + \Phi(2a)\Phi^2(b) + 2\Phi(a)\Phi(b)\Phi(a+b) - 4\Phi^2(a)\Phi^2(b),$$

$$\sigma_2 = -\Phi(a)\Phi(b+1)\Phi(a+b) - \Phi^2(a)\Phi(2b+1) + \Phi(a)\Phi(a+1)\Phi(2b) - \Phi(2a)\Phi(b)\Phi(b+1)$$
$$+\Phi(a+1)\Phi(b)\Phi(a+b) + \Phi^2(b)\Phi(2a+1) + 4\Phi^2(a)\Phi(b)\Phi(b+1) - 4\Phi(a)\Phi(a+1)\Phi^2(b),$$

$$\sigma_3^2 = \Phi^2(a)\Phi^2(b) \cdot \{\Phi(2a)\Phi^2(b+1) - 4\Phi^2(a)\Phi^2(b+1) + \Phi^2(a)\Phi(2(b+1))$$
$$+2\Phi(a)\Phi(b+1)\Phi(a+b+1) + \Phi^2(a+1)\Phi(2b) + \Phi^2(b)\Phi(2(a+1)) - 4\Phi^2(a+1)\Phi^2(b)$$
$$+2\Phi(a+1)\Phi(b)\Phi(a+b+1) - 2\Phi(a+1)\Phi(b+1)\Phi(a+b) - 2\Phi(a)\Phi(a+1)\Phi(2b+1)$$
$$+8\Phi(a)\Phi(b)\Phi(a+1)\Phi(b+1) - 2\Phi(2a+1)\Phi(b)\Phi(b+1) - 2\Phi(a)\Phi(b)\Phi(a+b+2)\}.$$

Consider the case of known β. The parameter Gamma can be represented in the form

$$\gamma = (\beta + a)\frac{\Phi(a)}{\Phi(a+1)},$$

the estimator is defined as

$$\gamma_n = (\beta + a)\frac{\Phi_n(a)}{\Phi_n(a+1)} \cdot \chi(\Phi_n(a+1) \geq \log^{-1} n)$$

and its asymptotic variance is equal to

$$\sigma_\gamma^2 = (\beta + a)^2 \frac{\Phi(2a)\Phi(a+1) + \Phi^2(a)\Phi(a+1)\Phi(2(a+1)) - 2\Phi(a)\Phi(2a+1)}{\Phi^3(a+1)} \tag{5}$$

Consider the optimization procedure of the parameter estimation of log-gamma distribution.

Define for an estimator γ_n of parameter γ the loss function

$$L_n = A(\gamma_n - \gamma)^2 + n.$$

Parameter A stands for a cost of mean square quality of the estimator γ_n of parameter γ and n is a sample size. We suppose that the cost of observations is included in the definition A (see, for comparison, [1]).

The corresponding risk function $R_n = EL_n$ has the form

$$R_n = AE(\gamma_n - \gamma)^2 + n$$

and we solve the optimization problem

$$R_n \to \min_n \tag{6}$$

Consider two cases.

– Case of known asymptotic variance σ^2 of γ_n.

Thus the principal term of the risk function has the form

$$R_n = \frac{A\sigma^2}{n} + n.$$

For A large enough the optimal sample size is equal to

$$n_A^0 = \sqrt{A\sigma^2}, \tag{7}$$

as well as the corresponding principal term of the risk function $R_{n_A^0}$

$$R_A^0 := \frac{A\sigma^2}{n_A^0} + n_A^0 = 2\sqrt{A\sigma^2}. \tag{8}$$

As follows the problem is solved if the number σ^2 is known.

– Case of unknown σ^2.

First define the estimator σ_n^2 of the variance σ^2 as

$$\sigma_n^2 = \frac{\Phi_n(2a)\Phi_n(a+1) + \Phi_n^2(a)\Phi^n(a+1)\Phi_n(2(a+1)) - 2\Phi_n(a)\Phi_n(2a+1)}{\Phi_n^3(a+1)} \cdot (b+a)^2\chi(\Phi_n(a+1) \geq \log^{-1} n). \tag{9}$$

Since (7) is directly involved in the expression (8) for R_A^0, the optimal sample size cannot be obtained as before. Similarly to Konev and Lai (1995), Sriram (1988), Sriram and Iaci (2014) and Kusainov and Vasiliev (2014), one uses the stopping time N_A as an estimator of n_A^0 replacing σ^2 in its definition with the estimator σ_n^2

$$N_A = \inf\{n \geq n_A : \ n \geq A^{1/2}\overline{\sigma}_A\}, \tag{10}$$

where $\overline{\sigma}_A = \min\{\sigma_{n_A}, \log A\}$, $\sigma_n = \sqrt{\sigma_n^2}$. We use here in comparison with mentioned above papers the estimator $\overline{\sigma}_A$ instead of σ_n to simplify the proofs. At the same time all results remain true.

It should be noted that for A large enough the following property is fulfilled

$$E(\overline{\sigma}_A^2 - \sigma^2)^{2p} \leq 2r_{n_A}(p), \tag{11}$$

where $r_n(p)$ is some deterministic sequence such that

$$A \cdot r_{n_A}(p) = o(1) \quad \text{as} \quad A \to \infty.$$

Indeed, for, e.g., $\log^2 A - \sigma^2 \geq 1$, using the Chebyshev inequality we have

$$E(\overline{\sigma}_A^2 - \sigma^2)^{2p} = E(\sigma_{n_A}^2 - \sigma^2)^{2p}\chi(\sigma_{n_A} \leq \log A) + (\log^2 A - \sigma^2)P(\sigma_{n_A} > \log A)$$

$$\leq r_{n_A}(p) + (\log^2 A - \sigma^2)\frac{E(\sigma_{n_A}^2 - \sigma^2)^{2p}}{(\log^2 A - \sigma^2)^{2p}} \leq 2r_{n_A}(p).$$

We prove the asymptotic equivalence of N_A and n_A^o in the almost surely and mean senses (see Theorem 1 below) and the optimality of the adaptive estimation procedure in the sense of equivalence of the obviously modified risk

$$\overline{R}_A := EL_{N_A} = AE(\gamma_{N_A} - \gamma)^2 + EN_A. \tag{12}$$

Theorem 1. *The observation numbers (10) and (7) and corresponding risk functions (12) and (8) are asymptotically equivalent in the following sense: as* $A \to \infty$

$$\frac{N_A}{n_A^o} \to 1 \quad a.s., \tag{13}$$

$$\frac{EN_A}{n_A^o} \to 1, \tag{14}$$

$$\frac{\overline{R}_A}{R_A^o} \to 1. \tag{15}$$

2 Simulation results

To illustrate the theoretical properties of the optimal adaptive procedure we give some numerical results for log-gamma distribution. We obtained the estimators σ_n^2 of the variance of parameter estimators. The results for different values of n are presented in Fig. 1. The horizontal line shows the asymptotic value of σ^2.

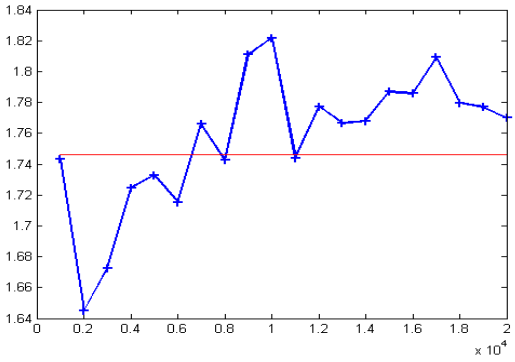


Figure 1: Log-gamma distribution. $\gamma = 1.666$

The quantities C_N and C_R are given in Fig. 2, 3 where $C_N = \frac{EN_A}{n_A^o}$, $C_R = \frac{\overline{R}_A}{R_A^o}$. Here n_A^0, R_A^0 are defined by (7, 8) and $N_A, \overline{R}_A$ – by (10, 12). Note that EN_A and $\overline{R}_A$ were computed as an empirical average over 1000 Monte Carlo replications of the experiment (for each value of A).

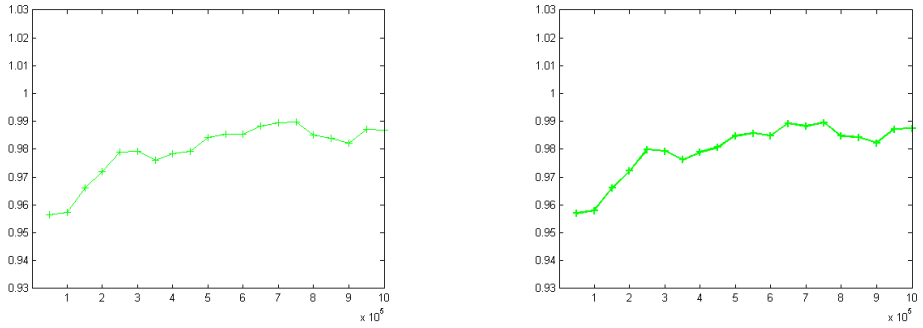


Figure 2: Log-gamma distribution. $\gamma = 1.6666$ C_N - left, C_R - right

The obtained numerical results are close to the theoretical properties of the proposed adaptive procedure.

Conclusion

The paper presents the method of optimal parameter estimation of log-gamma distribution. The truncated estimator is used to minimize the loss function which includes

the weighed mean square deviation and the sample size. It is shown that the proposed procedure is asymptotically efficient.

The theoretical results are illustrated by numerical results which confirm the optimality properties.

Acknowledgements

This study was supported by The Ministry of Education and Science of the Russian Federation, Goszadanie No 2.3208.2017/4.6

References

[1] Chernoff H. (1972). *Sequential Analysis and Optimal Design*. SIAM, Philadelphia.

[2] Dogadova T.V., Kusainov M.I., Vasiliev V.A. (2017). Truncated estimation method and applications. *Serdica. Mathematical Journal Bulgarian Academy of Sciences Institute of Mathematics and Informatics*. Vol. **43**, pp. 221-266.

[3] Konev V.V., Lai T L. (1995). Estimators with Prescribed Precision in Stochastic Regression Models. sl Sequential Analysis. Vol. **14**, pp. 179-192.

[4] Kusainov M.I., Vasiliev V.A. (2015). On optimal adaptive prediction of multivariate autoregression. *Sequential Analysis*. Vol. **34**, pp. 211-234.

[5] Politis D.N., Vasiliev V.A. (2012). Sequential kernel estimation of a multivariate regression function. *Proceedings of the IX International Conference 'System Identification and Control Problems, SICPRO'12, Moscow, 30 January – 2 February, 2012, V. A. Trapeznikov Institute of Control Sciences*. pp. 996-1009.

[6] Politis D.N., Vasiliev V.A. (2013). Non-parametric sequential estimation of a regression function based on dependent observations. *Sequential Analysis: Design Methods and Applications*. Vol. **32**, pp. 243–266. http://www.tandfonline.com/doi/full/10.1080/07474946.2013.803398. DOI:10.1080/07474946.2013.803398.

[7] Politis D.N., Vasiliev V.A., Vorobeychikov S. E. (2018). Truncated Estimation of Ratio Statistics with Application to Heavy Tail Distributions. *Math. Methods Statist*. Vol. **27**, pp. 226-243.

[8] Sriram T.N. (1988). Sequential Estimation of the Autoregressive Parameter in a First Order Autoregressive Process. *Sequential Analysis*. Vol. **7**, pp. 53-74.

[9] Sriram T.N., Ross , Iaci . (2014). Sequential Estimation for Time Series Models. *Sequential Analysis*. Vol. **33**, pp. 136-157.

[10] Vasiliev V. (2014) A truncated estimation method with guaranteed accuracy. *Annals of the Institute of Statistical Mathematics*. Vol. **66**. pp. 141–163.

Estimation of present value of deffered life annuity using information about expectation of life

Yury G. Dmitriev and Gennady M. Koshkin
National Research Tomsk State University, Tomsk, Russia
e-mail: dmit@mail.tsu.ru, kgm@mail.tsu.ru

Abstract

The paper deals with the estimation problem of the actuarial present value of the continuous deffered life annuity using auxiliary information about the expectation of life. Nonparametric estimators of life annuity are constructed by individuals' death moments. It is shown that the usage of such auxiliary information can often provide the mean squared error (MSE) smaller than that of standard estimators. An adaptive estimator is also proposed. The asymptotic normality of all these estimators is proved.

Keywords: nonparametric estimation; deffered life annuity; auxiliary information; mean squared error; asymptotic normality.

Introduction

Let x be the age of an individual and at the moment $t = 0$ payments start. The idea of the r-deferred life annuity in accordance with [4, p. 174] is this: from the moment $t+r = r$, an individual starts receiving money once a year, which we take as a monetary unit, and payments are made only during the lifetime of an individual. It is known that the deferred life annuity is associated with the appropriate type of insurance. Thus, the average total cost of the present continuous r-year deferred life annuity is given by the following formula (see [4, p. 184]):

$$ {}_{r|}\overline{a}_x(\delta) = \frac{1 - {}_{r|}\overline{A}_x}{\delta}, $$

where ${}_{r|}\overline{a}_x(\delta) = \int_r^\infty e^{-\delta t} f_x(t)dt$ is the net premium (the expectation of the present value of an insured unit sum for the deferred life insurance at age x), δ is a force of interest, $f_x(t) = \dfrac{f(x+t)}{S(x)}$ is a probability density of future lifetime $T(x) = X - x$ of an individual (x) [4, p. 62], $f(x)$ is a probability density of lifetime X of an individual (x), $S(x) = \mathbf{P}(X > x)$ is a survival function. Introduce the random variable

$$ z(x) = \frac{1 - e^{-\delta T(x)}}{\delta}, \quad T(x) > r. \tag{1} $$

Then, by averaging the random variable z(x) (1), we get the formula of the deferred life annuity (see [4]):

$$ {}_{r|}\overline{a}_x(\delta) = \mathbf{E}z(x) = \frac{1}{\delta}\left(1 - \frac{\Phi(x, \delta, r)}{S(x)}\right), \tag{2} $$

where $\mathbf{E}$ is the symbol of the mathematical expectation, $\Phi(x,\delta,r) = e^{\delta x}\int_{x+r}^{\infty} e^{-\delta t}dF(t)$, $F(x) = \mathbf{P}(X \le x) = 1 - S(x)$ is a distribution function.

Note that the whole life annuity $\overline{a}_x(\delta)$ is the special case of the deferred life annuity (2) at $r = 0$.

1 Construction of the Deferred Annuity Estimator

Assume that we have a random sample $X_1, \ldots, X_n$ of n individuals' lifetimes. Using the empirical survival function $S_n(x) = \dfrac{1}{n}\sum_{i=1}^{n} I(X_i > x)$, where $I(A)$ is the indicator of an event A, obtain the following estimator of (2):

$$ {}_{r|}\overline{a}_x^n(\delta) = \frac{1}{\delta}\left(1 - \frac{e^{\delta x}\sum_{i=1}^{n} e^{-\delta X_i} I(X_i > x+r)}{\sum_{i=1}^{n} I(X_i > x+r)}\right) = \frac{1}{\delta}\left(1 - \frac{\Phi_n(x,\delta,r)}{S_n(x)}\right), \qquad (3) $$

$$ \Phi_n(x,\delta,r) = \frac{e^{\delta x}}{n}\sum_{i=1}^{n} e^{-\delta X_i} I(X_i > x+r). $$

2 Bias and Mean Squared Error of Estimator ${}_{r|}\overline{a}_x^n(\delta)$

In this section, we will obtain the principal term of the asymptotic MSE and the bias convergence rate of the estimator (3). Introduce the notation according to [6]: the function $H(t) : R^s \to R^1$, where $t = t(x) = (t_1(x), \ldots, t_s(x))$ is an s-dimensional bounded function; $H_j(t) = \dfrac{\partial H(t)}{\partial t_j}$, $j = \overline{1,s}$, $\nabla H(t) = (H_1(t), \ldots, H_s(t))$; the symbol T denotes the transpose; $t_n = (t_{1n}, \ldots, t_{sn})$ is an s-dimensional statistic, $t_{jn} = t_{jn}(x) = t_{jn}(x, X_1, \ldots, X_n)$, $j = \overline{1,s}$; $\|t_n\| = \sqrt{t_{1n}^2 + \ldots + t_{sn}^2}$ is the Euclidean norm of t_n; $\Longrightarrow N_s\{\mu,\sigma\}$ is the symbol of weak convergence of sequence of random variables to the s-dimensional normal random variable with a mean $\mu = (\mu_1, \ldots, \mu_s)$ and symmetric covariance matrix $\sigma = ||\sigma_{ij}||$, $0 < \sigma_{jj} = \sigma_{jj}(x) < \infty$, $j = \overline{1,s}$; $\Re$ is the set of the integers.

D e f i n i t i o n. The function $H(t) : R^s \to R^1$ and the sequence $\{H(t_n)\}$ are said to belong to the class $\mathcal{N}_{\nu,s}(t;\gamma)$, provided that

1) there exists an ε-neighborhood $\{z : |z_i - t_i| < \varepsilon;\ i = \overline{1,s}\}$, in which the function $H(z)$ and all its partial derivatives $\dfrac{\partial H(z)}{\partial z_j}$ up to the order ν are continuous and bounded;

2) for any values of variables $X_1, ..., X_n$ the sequence $\{H(t_n)\}$ is dominated by a numerical sequence $C_0 d_n^{\gamma}$, such that $d_n \uparrow \infty$, as $n \to \infty$, and $0 \le \gamma < \infty$.

Theorem 1 [6]. Let the conditions 1) $H(z), \{H(t_n)\} \in \mathcal{N}_{2,s}(t,\gamma)$, 2) $\mathbf{E}||t_n - t||^i = O\left(d_n^{-i/2}\right)$ hold for all $i \in \Re$. Then, for every $k \in \Re$

$$\left|\mathbf{E}\left[H(t_n) - H(t)\right]^k - \mathbf{E}\left[\nabla H(t)(t_n - t)^T\right]^k\right| = O\left(d_n^{-(k+1)/2}\right). \tag{4}$$

If in formula (4) $k = 1$, we obtain the principal term $\mathbf{E}\left[\nabla H(t)(t_n - t)^T\right]$ of the bias $\mathbf{E}\left[H(t_n) - H(t)\right]$ for $H(t_n)$, and at $k = 2$, we have the principal term $\mathbf{E}\left[\nabla H(t)(t_n - t)^T\right]^2$ of the MSE $\mathbf{E}\left[H(t_n) - H(t)\right]^2$.

Denote $\tilde{C}({}_{r|}\overline{a}_x(\delta)) = \dfrac{\Phi(x, 2\delta, r)\, S(x) - \Phi^2(x, \delta, r)}{\delta^2\, S^3(x)}$.

Theorem 2 If the survival function $S(x) > 0$ and $S(t)$ is continuous at a point x, then

1) for the bias $b\left({}_{r|}\overline{a}_x^n(\delta)\right)$ of estimator (3) we have

$$\left|b\left({}_{r|}\overline{a}_x^n(\delta)\right)\right| = \left|\mathbf{E}\,{}_{r|}\overline{a}_x^n(\delta) - {}_{r|}\overline{a}_x(\delta)\right| = O\left(n^{-1}\right);$$

2) the MSE $u^2\left({}_{r|}\overline{a}_x^n(\delta)\right)$ is given by the formula

$$u^2({}_{r|}\overline{a}_x^n(\delta)) = \mathbf{E}\left({}_{r|}\overline{a}_x^n\delta) - {}_{r|}\overline{a}_x(\delta)\right)^2 = \frac{\tilde{C}({}_{r|}\overline{a}_x(\delta))}{n} + O\left(\frac{1}{n^{3/2}}\right).$$

Proof. For estimator ${}_{r|}\overline{a}_x^n(\delta)$ (3) in the notation of Theorem 1, we have: $s = 2$,

$$d_n = n,\ t_n = (t_{1n}, t_{2n}) = (\Phi_n(x,\delta,r),\ S_n(x)),\ H(t_n) = \frac{1}{\delta}\left(1 - \frac{\Phi_n(x,\delta,r)}{S_n(x)}\right) = {}_{r|}\overline{a}_x^n(\delta),$$

$$t = (t_1, t_2) = (\Phi(x,\delta,r),\ S(x)),\ H(t) = \frac{1}{\delta}\left(1 - \frac{t_1}{t_2}\right) = {}_{r|}\overline{a}_x(\delta),\quad H_1(t) = \frac{1}{\delta S(x)},$$

$$H_2(t) = -\frac{\Phi(x,\delta,r)}{\delta S^2(x)},\ \nabla H(t) = (H_1(t), H_2(t)) \neq 0.$$

Sequence $\{H(t_n)\}$ satisfies the condition 1) of Theorem 1 with $C_0 = \dfrac{1}{\delta}(1 + e^{-\delta r})$ and $\gamma = 0$. Indeed,

$$|H(t_n)| \le \frac{1}{\delta}\left(1 + \frac{\Phi_n(x,\delta,r)}{S_n(x)}\right) \le \frac{1}{\delta}\left(1 + \frac{e^{\delta x}\sum\limits_{i=1}^n e^{-\delta X_i} I(X_i > x + r)}{\sum\limits_{i=1}^n I(X_i > x + r)}\right) \le$$

$$\le \frac{1}{\delta}\left(1 + \frac{e^{\delta x}e^{-\delta(x+r)}\sum\limits_{i=1}^n I(X_i > x + r)}{\sum\limits_{i=1}^n I(X_i > x + r)}\right) = \frac{1}{\delta}(1 + e^{-\delta r}).$$

Further, in view of $t_2 = S(x) > 0$ the function $H(t)$ satisfies the condition 1) of Theorem 1. Also, this function satisfies the condition 2) of Theorem 1 due to Lemma 3.1 [5], as for all $i \in \Re$ such inequalities hold:

$$\mathbf{E}\{I^i\,(X > 0)\} = S(x) \le 1, \quad \mathbf{E}\{e^{i\delta\,x}e^{-i\delta\,X}I^i(X > 0)\} \le e^{i\delta\,x}e^{-i\delta\,x}S(x) = S(x) \le 1.$$

Therefore,

$$\mathbf{E}|\Phi_n(x,\delta,r) - \Phi(x,\delta,r)|^i = O\left(n^{-\frac{i}{2}}\right), \quad \mathbf{E}|S_n(x) - S(x)|^i = O\left(n^{-\frac{i}{2}}\right).$$

It is well known that $S_n(x)$ is the unbiased and consistent estimator of $S(x)$. Show that $\Phi_n(x,\delta,r)$ is the unbiased estimator of $\Phi(x,\delta,r)$:

$$\mathbf{E}\Phi_n(x,\delta,r) = \frac{e^{\delta x}}{n}\mathbf{E}\left\{\sum_{i=1}^{n} e^{-\delta\,X_i}I(X_i > x + r)\right\} = \Phi(x,\delta,r).$$

The ratio of two unbiased estimators can have a bias. Considering that all the conditions of Theorem 1 are fulfilled and $\mathbf{E}(t_n - t) = 0$, in accordance with (4) we get the order of the bias of ${}_{r|}\overline{a}_x^n(\delta)$:

$$\left|\mathbf{E}\left[{}_{r|}\overline{a}_x^n(\delta) - {}_{r|}\overline{a}_x(\delta)\right] - \mathbf{E}\left[\nabla H(t)(t_n - t)^T\right]\right| = \left|\mathbf{E}\left[{}_{r|}\overline{a}_x^n(\delta) - {}_{r|}\overline{a}_x(\delta)\right]\right| = O\left(n^{-1}\right).$$

Now, calculate the variance of $\Phi_n(x,\delta,r)$:

$$\mathbf{D}\Phi_n(x,\delta,r) = \frac{e^{2\delta x}}{n^2}\sum_{i=1}^{n}\mathbf{D}\left\{e^{-\delta\,X_i}I(X_i > x + r)\right\} = \frac{1}{n}\left(\Phi(x,2\delta,r) - \Phi^2(x,\delta,r)\right).$$

Similarly we find the components of covariance matrix $\sigma({}_{r|}\overline{a}_x(\delta)) = \begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{bmatrix}$ for statistics $\Phi_n(x,\delta,r)$ and $S_n(x)$:

$$\sigma_{11} = n\mathbf{D}\Phi_n(x,\delta,r) = \Phi(x,2\delta,r) - \Phi^2(x,\delta,r); \quad \sigma_{22} = n\mathbf{D}S_n(x) = S(x) - S^2(x);$$

$$\sigma_{12} = \sigma_{21} = n\mathrm{cov}\left(\Phi_n(x,\delta,r),\, S_n(x)\right) = (1 - S(x))\Phi(x,\delta,r).$$

Using the previous results on the bias and the covariance matrix, we obtain

$$u^2({}_{r|}\overline{a}_x^n(\delta)) = \mathbf{E}\left[\nabla H(t)(t_n - t)^T\right]^2 + O\left(\frac{1}{n^{3/2}}\right) =$$

$$= H_1^2(t)\sigma_{11} + H_2^2(t)\sigma_{22} + 2H_1(t)H_2(t)\sigma_{12} + O\left(\frac{1}{n^{3/2}}\right) = \frac{\tilde{C}({}_{r|}\overline{a}_x(\delta))}{n} + O\left(\frac{1}{n^{3/2}}\right). \quad (5)$$

The proof of Theorem 2 is completed.

3 Asymptotic Normality of Estimator ${}_{r|}\overline{a}_x^n(\delta)$

Theorem 3 (The usual central limit theorem) [1]. If $\xi_1, ..., \xi_n, ...$ is a sequence of independent and identically distributed s-dimensional vectors,

$$\mathbf{E}\xi_k = 0, \quad \sigma(x) = \mathbf{E}\{\xi_k^T \xi_k\}, \quad t_n = \frac{1}{n}\sum_{k=1}^{n} \xi_k,$$

then, as $n \to \infty$,

$$\sqrt{n} t_n \Longrightarrow N_s\{0, \sigma(x)\}.$$

Theorem 4 [6]. If $q_n(t_n - t) \Longrightarrow N_s\{\mu, \sigma\}$ for some number sequence $q_n \uparrow \infty$, the function $H(z)$ is differentiable at the point μ, $\nabla H(\mu) \neq 0$, then

$$q_n \left(H(t_n) - H(\mu)\right) \Longrightarrow N_1\{\nabla H(\mu)\, \mu^T,\, \nabla H(\mu)\, \sigma\, \nabla H^T(\mu)\}.$$

Theorem 5. Under the conditions of Theorem 2

$$\sqrt{n}[{}_{r|}\overline{a}_x^n(\delta) - {}_{r|}\overline{a}_x(\delta)] \Longrightarrow \mathsf{N}_1\left\{0, \tilde{C}({}_{r|}\overline{a}_x(\delta))\right\}.$$

Proof. In the notation of Theorem 3, we have $s = 2$, $\sigma(x) = \sigma({}_{r|}\overline{a}_x(\delta))$. Thus,

$$\sqrt{n}\,\{(\Phi_n(x, \delta, r),\, S_n(x)) - (\Phi(x, \delta, r),\, S(x))\} \Longrightarrow N_2\{(0, 0), \sigma({}_{r|}\overline{a}_x(\delta)\}.$$

The function $H(z)$ is differentiable at the point $t = (\Phi(x, \delta, r),\, S(x))$, $\nabla H(t) \neq 0$, and $q_n = \sqrt{n}$. Consequently, all the conditions of Theorem 4 hold, and using (5), we obtain the desired result.

The proof of Theorem 5 is completed.

4 Construction of Estimators Using Expected Lifetime

Suppose we know the expected lifetime

$$\mathbf{E}X = a. \tag{6}$$

The estimator by making use of such information according to [2, 3] can be taken in the following form:

$${}_{r|}\overline{a}_x^n(\delta, \lambda) = \frac{1}{\delta}\left(1 - \frac{\Phi_n(x, \delta, r)}{S_n(x)} - \lambda(\overline{x} - a)\right), \tag{7}$$

where $\overline{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ is an estimator of a, parameter λ we will find minimizing the principal term of the asymptotic MSE of ${}_{r|}\overline{a}_x^n(\delta, \lambda)$ (7). Estimator (7) combines the available empirical information containing in (3) and prior information (6).

For estimator ${}_{r|}\overline{a}_x^n(\delta,\lambda)$ in the notation of Theorem 1, we have: $s=3$; $d_n=n$;

$t_n=(t_{1n},t_{2n},t_{3n})=(\Phi_n(x,\delta,r),S_n(x),\overline{x})$; $t=(t_1,t_2,t_3)=(\Phi(x,\delta,r),S(x),a)$;

$$H(t_1,t_2,t_3)=\frac{1}{\delta}\left(1-\frac{t_1}{t_2}-\lambda(t_3-a)\right)=\frac{1}{\delta}\left(1-\frac{\Phi(x,\delta,r)}{S(x)}-\lambda(a-a)\right)={}_{r|}\overline{a}_x(\delta,\lambda),$$

$$H(t_n)=\frac{1}{\delta}\left(1-\frac{\Phi_n(x,\delta,r)}{S_n(x)}-\lambda(\overline{x}-a)\right)={}_{r|}\overline{a}_x^n(\delta,\lambda),$$

$$\nabla H(t)=(H_1(t),H_2(t),H_3(t))=\left(\frac{1}{\delta S(x)},-\frac{\Phi(x,\delta,r)}{\delta S^2(x)},-\frac{\lambda}{\delta}\right)\neq 0. \tag{8}$$

5 Bias and MSE of ${}_{r|}\overline{a}_x^n(\delta,\lambda)$

Arguing as in Section 2, it is easy to show that the sequence $\{H(t_n)\}$ satisfies the condition 1) of Theorem 1 with $C_0=\frac{1}{\delta}(1+e^{-\delta r}+|\lambda|(\omega+a))$, where $\omega<\infty$ is the limiting age and $\gamma=0$; also, the statistic t_n satisfies the condition 2) due to Lemma 3.1 [5], provided that $\mathbf{E}X^i\leq\omega^i<\infty$ for all $i\in\Re$. Hence, given that $\mathbf{E}(t_n-t)=0$, for the bias of (7) we obtain the following result:

$$\left|\mathbf{E}\left[{}_{r|}\overline{a}_x^n(\delta,\lambda)-{}_{r|}\overline{a}_x(\delta)\right]\right|=\left|b\left[{}_{r|}\overline{a}_x^n(\delta,\lambda)\right]\right|=O\left(n^{-1}\right). \tag{9}$$

Now, find the covariance matrix $\sigma({}_{r|}\overline{a}_x(\delta,\lambda))=\begin{bmatrix}\sigma_{11}&\sigma_{12}&\sigma_{13}\\ \sigma_{21}&\sigma_{22}&\sigma_{23}\\ \sigma_{31}&\sigma_{32}&\sigma_{33}\end{bmatrix}$ for statistics $\Phi_n(x,\delta,r)$, $S_n(x)$, and $\overline{x}$: $\sigma_{23}=\sigma_{32}=n\mathrm{cov}\left(S_n(x),\,\overline{x}\right)=C_2(x,r)-aS(x)$;

$$\sigma_{13}=\sigma_{31}=n\mathrm{cov}\left(\Phi_n(x,\delta,r),\,\overline{x}\right)=C_1(x,\delta,r)-a\Phi(x,2\delta,r);\quad \sigma_{33}=n\mathbf{D}\overline{x}=\mathbf{D}X,$$

where $C_1(x,\delta,r)=e^{\delta x}\int_{x+r}^{\infty}e^{-\delta u}udF(u)$, $C_2(x,r)=\int_{x+r}^{\infty}udF(u)$, and σ_{11}, σ_{22}, $\sigma_{12}=\sigma_{21}$ are defined in Section 2. Using (5) and (8), the above results for bias (9) and covariance matrix $\sigma({}_{r|}\overline{a}_x(\delta,\lambda))$, we obtain:

$$u^2({}_{r|}\overline{a}_x^n(\delta,\lambda))=\mathbf{E}\left[\nabla H(t)(t_n-t)^T\right]^2+O\left(\frac{1}{n^{3/2}}\right)=\frac{\tilde{C}({}_{r|}\overline{a}_x(\delta,\lambda))}{n}+O\left(\frac{1}{n^{3/2}}\right), \tag{10}$$

$$\tilde{C}({}_{r|}\overline{a}_x(\delta,\lambda))=\sum_{p=1}^{3}\sum_{j=1}^{3}H_p(t)\sigma_{pj}H_j(t)=\tilde{C}({}_{r|}\overline{a}_x(\delta))+\lambda^2Q_1-2\lambda Q_2,$$

where $Q_1=\frac{\sigma_{33}}{\delta^2}>0$, $Q_2=\frac{H_1(t)\sigma_{13}+H_2(t)\sigma_{23}}{\delta}$. The minimum of $\tilde{C}({}_{r|}\overline{a}_x(\delta,\lambda))$ with respect to λ is achived at $\lambda_0=\frac{Q_2}{Q_1}$. Such λ_0 minimizes the principal term of MSE (10), and this minimum is as follows:

$$u^2({}_{r|}\overline{a}_x^n(\delta,\lambda_0))=\frac{\tilde{C}({}_{r|}\overline{a}_x(\delta,\lambda_0))}{n}+O\left(\frac{1}{n^{3/2}}\right)=\frac{1}{n}\left(\tilde{C}({}_{r|}\overline{a}_x(\delta))-\frac{Q_2^2}{Q_1}\right)+O\left(\frac{1}{n^{3/2}}\right).$$

So, the constant of the principal term of MSE (10) is less than the constant of the principal term of MSE (5), i.e.,

$$\tilde{C}({}_{r|}\overline{a}_x(\delta,\lambda_0)) = \tilde{C}({}_{r|}\overline{a}_x(\delta)) - \frac{Q_2^2}{Q_1} < \tilde{C}({}_{r|}\overline{a}_x(\delta)). \tag{11}$$

In accordance with (11), the estimator

$${}_{r|}\overline{a}_x^n(\delta,\lambda_0) = \frac{1}{\delta}\left(1 - \frac{\Phi_n(x,\delta,r)}{S_n(x)} - \lambda_0(\overline{x}-a)\right), \tag{12}$$

will be called the optimal (in the mean square sense) estimator. The non-negative quantity $\dfrac{Q_2^2}{Q_1}$ in (11) determines the decrease of the principal term of MSE for the optimal estimator by using auxiliary information (6).

6 Adaptive Estimator

Theorem 6. For the optimal estimator ${}_{r|}\overline{a}_x^n(\delta,\lambda_0)$ under the conditions of Theorem 2

$$\sqrt{n}\left[{}_{r|}\overline{a}_x^n(\delta,\lambda_0) - {}_{r|}\overline{a}_x(\delta)\right] \Longrightarrow \mathsf{N}_1\left\{0, \tilde{C}({}_{r|}\overline{a}_x(\delta,\lambda_0))\right\}.$$

Proof. The statement of Theorems 6 follows from Theorem 4 with the usage of the arguments of Sections 3–5.

The statistic ${}_{r|}\overline{a}_x^n(\delta,\lambda_0)$ can be used as an estimator for ${}_{r|}\overline{a}_x(\delta)$ if we know λ_0; otherwise, it is required to construct an adaptive estimator. We need a more detailed formula for λ_0:

$$\lambda_0 = \frac{1}{S(x)\mathbf{D}X}\left(\frac{\Phi(x,\delta,r)}{S(x)}\left\{C_2(x,r) - aS(x)\right\} - C_1(x,\delta,r) + a\Phi(x,\delta,r)\right). \tag{13}$$

According to (13) construct the estimator

$$\hat{\lambda}_0 = \frac{1}{s^2 S_n(x)}\left(\frac{\Phi_n(x,\delta,r)}{S_n(x)}\left\{\hat{C}_2(x,r) - aS_n(x)\right\} - \hat{C}_1(x,\delta,r) + a\Phi_n(x,\delta,r)\right), \tag{14}$$

where $s^2 = \dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{x})^2$ is an unbiased estimator of the variance $\mathbf{D}X$,

$$\hat{C}_2(x,r) = n^{-1}\sum_{i=1}^{n} X_i I(X_i > x+r),\ \hat{C}_1(x,\delta,r) = n^{-1}\sum_{i=1}^{n} e^{-\delta X_i} X_i I(X_i > x+r).$$

Theorem 7. For the adaptive estimator ${}_{r|}\overline{a}_x^n(\delta,\hat{\lambda}_0)$ under the conditions of Theorem 2

$$\sqrt{n}\left[{}_{r|}\overline{a}_x^n(\delta,\hat{\lambda}_0) - {}_{r|}\overline{a}_x(\delta)\right] \Longrightarrow \mathsf{N}_1\left\{0, \tilde{C}({}_{r|}\overline{a}_x(\delta,\lambda_0))\right\}.$$

Proof. The following equality holds:

$$\sqrt{n}\left[{}_{r|}\overline{a}_x^n(\delta,\hat{\lambda}_0)-{}_{r|}\overline{a}_x(\delta)\right]=\sqrt{n}\left[{}_{r|}\overline{a}_x^n(\delta,\lambda_0)-{}_{r|}\overline{a}_x(\delta)\right]+R_n,$$

where $R_n=\delta^{-1}(\lambda_0-\hat{\lambda}_0)\sqrt{n}(\overline{x}-a)$. All the estimators, used in (14), converge almost surely to their true values according to the strong law of large numbers (the Second Theorem of Kolmogorov [7]. Thus, from the First Continuity Theorem of Borovkov [1], estimator $\hat{\lambda}_0$ converges almost surely to λ_0. Based on the central limit theorem $\sqrt{n}(\overline{x}-a)\Longrightarrow \mathsf{N}_1\{0,\mathbf{D}X\}$, we retrieve $R_n\Longrightarrow 0$. Now, the statement of Theorem 7 is proved by making use of Theorem 6.

Conclusions

The paper deals with the estimation problem of the present values of the continuous deffered life annuity using auxiliary information about the expectation of life. It is shown that the usage of such auxiliary information can often provide the MSE smaller than that of standard estimators. We proved the results on asymptotic properties of the proposed estimators: unbiasedness, consistency and normality. Also, the principal terms of the asymptotic MSEs of the estimators were found. An adaptive estimator is constructed; such estimator is equivalent (in the sense of asymptotic distribution) to the estimator with the optimal weight coefficient λ_0.

References

[1] Borovkov A.A. (1998) *Mathematical Statistics*. Gordon & Breach, New York.

[2] Dmitriev Yu.G., Koshkin G.M. (1987). On the Use of a Priori Information in Monparametric Regression Estimation. *IFAC Proceedings Series*. Vol. **2**, pp. 223-228.

[3] Dmitriev Yu.G., Koshkin G.M. (2018). Nonparametric Estimators of Probability Characteristics Using Unbiased Prior Conditions. *Statistical Papers*. Vol. **59**. No. **4**, pp. 1559-1575.

[4] Falin G.I. (2002). *Mathematical Foundations of the Theory of Life Insurance and Pension Schemes*. Ankil Publ., Moscow (in Russian).

[5] Ibragimov I.A., Khasminskii R.Z. (1981). *Statistical Estimation: Asymptotic Theory*. Springer, Berlin; New York.

[6] Koshkin G.M. (1999). Deviation Moments of the Substitution Estimator and Its Piecewise Smooth Approximations. *Siberian Mathematical Journal*. Vol. **40**. No. **3**, pp. 515-527.

[7] Rao C.R. (1965). *Linear Statistical Methods and Its Applications*. Willey, New York.

Robust extrapolation in discrete systems with random jump parameters and incomplete information

Valeriy I. Smagin, Gennady M. Koshkin and Konstantin S. Kim
National Research Tomsk State University, Tomsk, Russia
e-mail: vsm@mail.tsu.ru, kgm@mail.tsu.ru, kks93@rambler.ru

Abstract

An algorithm for the synthesis of a robust extrapolator is considered, which determines the estimate of the state vector of a discrete linear system with random jump parameters described by a Markov chain with a finite number of states under incomplete information about the model and the observation channel. The transfer matrix of the extrapolator are invited to choose independent of the state of the jump process using the nonparametric smoothing procedure.

Keywords: discrete model, robust extrapolation, jump parameters, incomplete information, nonparametric smoothing.

Introduction

The problem of constructing estimates of extrapolation and filtering under incomplete information were considered in [3–6, 12, 13, 15]. In these papers, problems of estimating using recurrent algorithms of the Kalman type under the condition of the presence of unknown inputs in the model were considered, and the problem of estimation in the presence of an unknown vector in the observation channel was also considered in [16]. In [1, 2, 9–11, 14], estimation problems in systems with random jump parameters were considered. Such problems arise when building models of real processes with possible failures. In this paper, we consider the problem of synthesizing a robust extrapolator for a discrete object with random jump parameters with a finite number of states. Problems are considered for objects and observation channels with unknown parameters and unknown additive vectors. Using the procedure of non-parametric smoothing, the problem of synthesizing a robust extrapolator is solved, the transfer coefficients of which do not depend on the state of the jump process. The simulation results are given.

1 Statement Problem

We consider the discrete stochastic system, which is described by the equation

$$x(k+1) = (A_\gamma + \Delta A_\gamma)x(k) + B_\gamma f(k) + q_\gamma(k),\ x(0) = x_0, \tag{1}$$

and available observations are set as follows:

$$y(k) = (S_\gamma + \Delta S_\gamma)x(k) + H_\gamma \phi(k) + v_\gamma(k), \tag{2}$$

where $x(k) \in R^m$ is a state of the system; $\gamma = \gamma(k)$ is a jump parameter (Markov chain with n states $\gamma_1, \gamma_2, ..., \gamma_n$); $f(k)$, $\phi(k)$ are unknown vectors; x_0 is a random vector ($\bar{x}_0 = \mathsf{E}\{x_0\}$ and $N_{0,i} = \mathsf{E}\{(x_0 - \bar{x}_0) \times (x_0 - \bar{x}_0)^T / \gamma = \gamma_i\}, i = \overline{1,n}$); $y(k) \in R^l$ is the observation vector; A_γ, B_γ, S_γ, H_γ are given matrices; ΔA_γ, ΔS_γ are unknown matrices; $q_\gamma(k)$, $v_\gamma(k)$ are independent random sequences with the following characteristics: $\mathsf{E}\{q_\gamma(k)\} = 0$, $\mathsf{E}\{v_\gamma(k)\} = 0$, $\mathsf{E}\left\{q_\gamma(k) q_\gamma^T(j)\right\} = Q_\gamma \delta_{kj}$, $\mathsf{E}\left\{v_\gamma(k) v_\gamma^T(j)\right\} = V_\gamma \delta_{kj}$ (E denotes expectation and T denotes transposition of a matrix, δ_{kj} is the Kronecker symbol).

The probability $p_i(k) = P\{\gamma(k) = \gamma_i\}$, $i = \overline{1,n}$, satisfies the equation

$$p_i(k+1) = \sum_{j=1}^{n} p_{ij} p_j(k),\ p_i(0) = p_{i,0}, \tag{3}$$

where p_{ij} is the probability of transition from the state i to the state j in one step, $p_{i,0}$ is the initial probability of the i-th state. According to the information received at the k-th step, it is required to find estimates of extrapolation $\widehat{x}(k+1)$ based on the minimization of the following criterion:

$$\begin{aligned} J[0;T_f] = \frac{1}{T_f} \mathsf{E}\{[\sum_{k=0}^{T_f} \sum_{i=1}^{n} p_i(k) e^T(k) R_i(k) e(k) + \\ + \sum_{i=1}^{n} p_i(T_f) e^T(T_f) L_i(T_f) e(T_f)] / \gamma(0) = \gamma_0\}, \end{aligned} \tag{4}$$

where $e(k) = x(k) - \widehat{x}(k)$, $R_i(k) > 0$, $L_i(k) > 0$ are weight matrices, γ_0 is the initial value of jump parameter γ.

2 Optimization of the Criterion

We present the system (1) in the following form:

$$x(k+1) = A_\gamma x(k) + r(k) + q_\gamma(k),\ x(0) = x_0, \tag{5}$$

where $r(k) = \Delta A_\gamma x(k) + B_\gamma f(k)$ is an unknown input. The channel of observations we will present as

$$y(k) = S_\gamma x(k) + \varphi(k) + v_\gamma(k), \tag{6}$$

where $\varphi(k) = \Delta S_\gamma x(k) + H_\gamma \phi(k)$ is unknown vector in the observation channel.

To solve the problem of synthesizing a robust extrapolator, we use the separation principle. This means that we first construct estimates of the vector under the assumption that the vectors $r(k)$ and $\varphi(k)$ are known. For this, we use the recurrent algorithm of Kalman extrapolation [1]

$$\widehat{x}(k+1) = A_\gamma \widehat{x}(k) + r(k) + K(k)(y(k) - S_\gamma \widehat{x}(k) - \varphi(k)),\ \widehat{x}(0) = \bar{x}_0. \tag{7}$$

In (7), we will look for a matrix $K(k)$ independent of $\gamma(k)$ that will ensure the extrapolator robustness with respect to the error of jump parameter $\gamma(k)$. Then, the

estimates of the vectors $\widehat{r}(k)$ and $\widehat{\varphi}(k)$ are constructed under the assumption that the prediction estimate of the state vector $\widehat{x}(k)$ is known.

We introduce the notation for the matrices $Q_\gamma, V_\gamma, R_\gamma, N_\gamma, L_\gamma, A_\gamma, S_\gamma$ at $\gamma = \gamma_i$ as $Q_i, V_i, R_i, N_i, L_i, A_i, S_i$, respectively ($i = \overline{1,n}$). Consider a theorem, in which we construct an algorithm determining the matrix $K(k)$ for extrapolator (7) based on optimization of criterion (4).

Theorem. Let there exist positive definite matrices N_i and L_i, which are the solution of a two-point boundary value problem:

$$N_i(k+1) = (A_i - K(k)S_i)(\sum_{j=1}^{n} p_{i,j}N_j(k))(A_i - K(k)S_i)^T + Q_i + \\ + K(k)V_iK(k)^T,\ N_i(0) = N_0, \tag{8}$$

$$L_i(k) = (A_i - K(k)S_i)^T(\sum_{j=1}^{n} p_{i,j}L_j(k+1))(A_i - K(k)S_i)^T + R_i,\ L_i(T_f) = L_{T,i}. \tag{9}$$

Then, the vector $\mathrm{ct}(K(k))$, composed of transpose rows of the matrix $K(k)$ and providing the minimum of criterion (4), is determined by the formula:

$$\mathrm{ct}(K(k)) = (\sum_{i=1}^{n} p_i(k+1)[L_i(k+1) \otimes S_i(\sum_{j=1}^{n} p_{i,j}N_j(k))S_i^T + \\ + L_i(k+1) \otimes V_i])^{-1}\mathrm{ct}(\sum_{i=1}^{n} p_i(k+1)L_i(k+1)A_i(\sum_{j=1}^{n} p_{i,j}N_j(k)S_i^T)). \tag{10}$$

In (10), the symbol $\otimes$ denotes the Kronecker product.

Proof. Let's present the criterion (4) as a sum

$$J[0;T_f] = \sum_{i=1}^{n} J_i[k,T_f],\ k = \overline{0,T_f}. \tag{11}$$

In (11) $J_i[k,T_f]$ is determined by the formula

$$J_i[k;T_f] = \sum_{\xi=k}^{T_f-1} \mathrm{tr}\, p_i(\xi)N_i(\xi)R_i(\xi) + \mathrm{tr}\, p_i(T_f)N_i(T_f)L_i(T_f), \tag{12}$$

where tr is the trace of a matrix.

Introduce the following Lyapunov function:

$$W(k, N_i(k)) = \mathrm{tr}\, p_i(k)N_i(k)R_i(k) + \mathrm{tr} \sum_{t=k}^{T_f} p_i(t)\bar{\Psi}_i(t)L_i(t), \tag{13}$$

where $L_i(t)$ is determined by equation (9), $\bar{\Psi}_i(t) = Q_i + K(t)V_iK(t)^T + \Psi_i(t)$ ($\Psi_i(t)$ is some positive definite matrix).

Summing over $k = \overline{t, T_f - 1}$, the finite differences of the function $W(k, N_i(k))$, taking into account formula (9), we obtain:

$$\sum_{k=t}^{T_f} \Delta W(k, N_I(k)) = \sum_{k=t}^{T_f-1} [W(k+1, N_i(k+1) - W(k, N_i(k))] = \\ = \sum_{k=t}^{T_f-1} \mathrm{tr}[p_i(k+1)N_i(k+1)L_i(k+1) - p_i(k)N_i(k)L_i(k) - p_i(k)\bar{\Psi}_i(k)L_i(k)]. \tag{14}$$

On the other hand, this expression can be represented as

$$\sum_{k=t}^{T_f} \Delta W(k, N_i(k)) = W(t+1, N_i(t+1)) - W(t, N_i(t)) + ...+ \\ + W(T_f, N_i(T_f)) - W(T_f - 1, N_i(T_f - 1) = \\ = \mathrm{tr}\, p_i(T_f)N_i(T_f)L_i(T_f) - \mathrm{tr}\, p_i(t)N_i(t)L_i(t) - \mathrm{tr} \sum_{\xi=t}^{T_f-1} p_i(\xi)\bar{\Psi}_i(\xi)L_i(\xi). \tag{15}$$

Add into (12) the difference of the right parts of (14) and (15). Considering that this difference is zero, criterion (11) takes the form:

$$J[0; T_f] = \sum_{i=1}^{n} \{ \sum_{\xi=k}^{T_f-1} \mathrm{tr} p_i(\xi)N_i(\xi)R_i(\xi) - \sum_{\xi=k+1}^{T_f-1} \mathrm{tr} p_i(\xi)N_i(\xi)L_i(\xi)+ \\ + \sum_{\xi=k}^{T_f-1} \mathrm{tr}\, p_i(\xi+1)[(A_i - K(\xi)S_i)(\sum_{j=1}^{n} p_{i,j}N_j(\xi))(A_i - K(\xi)S_i)^T + \\ + Q_i + K(\xi)V_iK(\xi)^T]L_i(\xi+1)\}. \tag{16}$$

Now, calculate the derivatives:

$$\frac{dJ}{dK(k)} = \sum_{\xi=k}^{T_f-1} \sum_{i=1}^{n} \mathrm{tr}[-L_i(\xi+1)p_i(\xi+1)A_i(\sum_{j=1}^{n} p_{i,j}N_j(\xi))S_i^T - \\ -p_i(\xi+1)L_i(\xi+1)A_i(\sum_{j=1}^{n} p_{i,j}N_j(\xi))S_i^T + p_i(\xi+1)L_i(\xi+1)K(\xi)S_i \times \\ \times(\sum_{j=1}^{n} p_{i,j}N_j(\xi))S_i^T + L_i(\xi+1)p_i(\xi+1)K(\xi)S_i(\sum_{j=1}^{n} p_{i,j}N_j(\xi))S_i^T + \\ +p_i(\xi+1)L_i(\xi+1)K(\xi)V_i + L_i(\xi+1)p_i(\xi+1)K(\xi)V_i]. \tag{17}$$

Equating this derivative to zero, assuming that each summand of summation with respect to i is zero, and using the Kronecker product operation [7], we obtain formula (10) determining the matrix $K(k)$.

Calculate the finite difference of the Lyapunov function:

$$\begin{gathered}\Delta W(k,N_i(k)=W(k+1,N_i(k+1))-W(k,N_i(k))=\\ =\mathrm{tr}p_i(k+1)N_i(k+1)R_i(k+1)+\\ +\mathrm{tr}\sum_{t=k+1}^{T_f}p_i(t)[Q_i+K(t)V_iK(k)^T+\Psi_i(t)]L_i(t)-\\ -\mathrm{tr}p_i(k)N_i(k)R_i(k)-\mathrm{tr}\sum_{t=k}^{T_f}p_i(t)[Q_i+K(t)V_iK(k)^T+\Psi_i(t)]L_i(t)=\\ =\mathrm{tr}p_i(k+1)N_i(k+1)R_i(k+1)-\mathrm{tr}p_i(k)N_i(k)R_i(k)-\\ -p_i(k)[Q_i+K(k)V_iK(k)^T+\Psi_i(k)]L_i(k).\end{gathered}\tag{18}$$

Since the matrices $N_i,\ L_i$ are positively determined by Theorem conditions, and the matrix $\Psi_i(t)>0$ is given arbitrarily, it is obvious that these matrices can be chosen such that the final difference (18) becomes negative. This condition guarantees the Lyapunov stability of the extrapolator dynamic. Theorem is proved.

To construct the prediction estimate, we use the Kalman extrapolator

$$\widehat{x}(k+1)=A_\gamma\widehat{x}(k)+\widehat{r}(k)+K(k)(y(k)-S_\gamma\widehat{x}(k)-\widehat{\varphi}(k)),\ \widehat{x}(0)=\bar{x}_0,\tag{19}$$

where $K(k)$ is the transfer matrix depending on k and independing of jump parameter $\gamma(k)$.

3 Synthesis of the Stationary Extrapolator

In this case, the matrix of transfer coefficients $K(k)$ in Kalman extrapolator (19) will be constant, and the criterion will take the form:

$$J[0;\infty]=\lim_{T_f\to\infty}\frac{1}{T_f}\sum_{k=1}^{T_f}\sum_{i=1}^{n}\mathrm{tr}\bar{p}_iN_i(k)R_i(k).\tag{20}$$

The two-point boundary value problem is transformed into the following matrix equations:

$$N_i=(A_i-KS_i)(\sum_{j=1}^{n}p_{i,j}N_j)(A_i-KS_i)^T+Q_i+KV_iK^T,\tag{21}$$

$$L_i=(A_i-KS_i)^T(\sum_{j=1}^{n}p_{i,j}L_j)(A_i-KS_i)^T+R_i,\tag{22}$$

$$\mathrm{ct}(K)=(\sum_{i=1}^{n}\bar{p}_i[L_i\otimes S_i(\sum_{j=1}^{n}p_{i,j}N_j)S_i^T+L_i\otimes V_i])^{-1}\mathrm{ct}(\sum_{i=1}^{n}\bar{p}_iL_iA_i(\sum_{j=1}^{n}p_{i,j}N_jS_i^T)),\tag{23}$$

where $\bar{p}_i$ are steady-state probabilities (solution of system (3) as $k\to\infty$).

Thus, to synthesize a stationary extrapolator, it is necessary to solve the system of matrix equations (21)–(23).

Note that if there are positively defined solutions N_i, L_i $(i = \overline{1,n})$ of matrix equations (21)–(23), then using the equation (22) and condition $R_i > 0$ (see Theorem 1.6 [8]), it follows that a stationary extrapolator with jump parameters will be stochastic stable.

4 Estimates of Unknown Vectors

As estimates of unknown vectors, various algorithms can be used [3–5]. When using the LSM estimates, finding $\widehat{\varphi}(k)$ and $\widehat{r}(k)$ is based on minimizing of the following criteria:

$$J_1 = \sum_{i=1}^{k}(\left\|y(t) - S_t\widehat{x}(t)\right\|_{W_1}^2 + \left\|\varphi(t-1)\right\|_{\bar{W}_1}^2), \tag{24}$$

$$J_2 = \sum_{i=1}^{k}(\left\|y(t) - \widehat{\varphi}(t) - S_\gamma A_\gamma\widehat{x}(t-1)\right\|_{W_2}^2 + \left\|r(t-1)\right\|_{\bar{W}_2}^2), \tag{25}$$

where W_1, $\bar{W}_1$, W_2, $\bar{W}_2$ are weight matrices. At each step, the estimates of $\widehat{\varphi}(k)$ and $\widehat{r}(k)$ are constructed sequentially, first minimizing the criterion (24), then (25). In constructing vector estimates $\widehat{r}(k)$, based on the criterion (25), are used $\widehat{\varphi}(k)$ vector estimates obtained via the minimization of the criterion (24). Then, estimates of unknown vectors by the LSM estimates are determined as follows:

$$\widehat{\varphi}^{\langle LSM\rangle}(k) = [S_\gamma^T W_1 S_\gamma + \bar{W}_1]^{-1} S_\gamma^T W_1\{y(k) - S_\gamma\widehat{x}(k)\}, \tag{26}$$

$$\widehat{r}^{\langle LSM\rangle}(k) = [S_\gamma^T W_2 S_\gamma + \bar{W}_2]^{-1} S_\gamma^T W_2\{y(k) - \widehat{\varphi}(k) - S_\gamma A_\gamma\widehat{x}(k-1)\}. \tag{27}$$

By analogy with [6], using estimates (26) and (27), we construct prediction estimates by making use of the technique of nonparametric smoothing:

$$\widehat{\varphi}^{\langle NP\rangle}(k) = [S_\gamma^T W_1 S_\gamma + \bar{W}_1]^{-1} S_\gamma^T W_1\Omega(k), \tag{28}$$

$$\widehat{r}^{\langle NP\rangle}(k) = [S_\gamma^T W_2 S_\gamma + \bar{W}_2]^{-1} S_\gamma^T W_2\bar{\Omega}(k). \tag{29}$$

In (28) and (29) the components of the vectors Ω and $\bar{\Omega}$ are determined by the formulas:

$$\Omega_j(k) = \frac{\sum_{i=1}^{k}\frac{[y(i)-S_\gamma(\widehat{x}(i))]_j}{\mu_j}G(\frac{k-i+1}{\mu_j})}{\sum_{i=1}^{k}\frac{1}{\mu_j}G(\frac{k-i+1}{\mu_j})}\ (j = \overline{1,n}), \tag{30}$$

$$\bar{\Omega}_s(k) = \frac{\sum_{i=1}^{k}\frac{[y(i)-\widehat{\varphi}(i)-S_\gamma A_\gamma(\widehat{x}(i-1))]_s}{\bar{\mu}_s}G(\frac{k-i+1}{\bar{\mu}_s})}{\sum_{i=1}^{k}\frac{1}{\bar{\mu}_s}G(\frac{k-i+1}{\bar{\mu}_s})}\ (s = \overline{1,n}), \tag{31}$$

where in relations (30), (31) $G(\cdot)$ is a kernel function, μ_j and $\bar{\mu}_s$ are bandwidth parameters.

5 Simulation Results

The simulation was performed for the following data:

$$A_1 = \begin{pmatrix} 0.85 & 0.1 \\ -0.05 & 0.94 \end{pmatrix}, A_2 = \begin{pmatrix} 0.89 & 0.05 \\ -0.02 & 0.45 \end{pmatrix}, B_1 = B_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$Q_1 = Q_2 = \begin{pmatrix} 0.05 & 0 \\ 0 & 0.03 \end{pmatrix}, \Delta A_1 = \begin{pmatrix} 0.05 & 0.01 \\ 0.02 & -0.11 \end{pmatrix}, \Delta A_2 = \begin{pmatrix} 0.06 & 0.012 \\ -0.01 & 0.05 \end{pmatrix},$$

$$S_1 = S_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \Delta S_1 = \begin{pmatrix} 0.01 & 0 \\ 0 & 0 \end{pmatrix}, \Delta S_2 = \begin{pmatrix} 0.008 & 0 \\ 0 & 0 \end{pmatrix},$$

$$V_1 = V_2 = \begin{pmatrix} 0.01 & 0 \\ 0 & 0.15 \end{pmatrix}, H_1 = H_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, P = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix},$$

$$R_1 = R_2 = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.15 \end{pmatrix}, W_1 = W_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \bar{W}_1 = \bar{W}_2 = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix},$$

$$f(k) = \begin{pmatrix} 0.1 + 0.1\sin(0.1k) \\ 0.1 + 0.12\sin(0.15k) \end{pmatrix}, \phi(k) = \begin{pmatrix} 0.3 \\ 0.25 \end{pmatrix}.$$

We use the Gaussian kernels

$$G(u) = \frac{exp(\frac{-u^2}{2})}{\sqrt{2\pi}}.$$

Fig.1 shows the graphs of the values of the jump parameter γ.

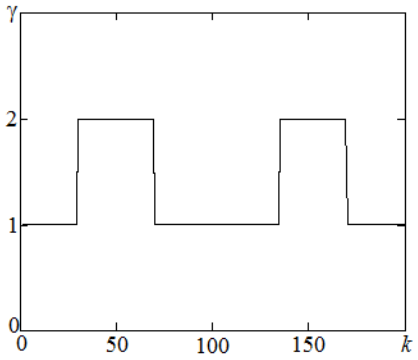


Figure 1: The Values of the Jump Parameter γ

The transfer matrix K of the extrapolator is determined from the solution of matrix equations (21)–(23).

The root-mean-square errors were calculated as follows:

$$\sigma_i(N) = \sqrt{\frac{\sum_{k=1}^{N}(x_i(k) - \widehat{x}_i(k))^2}{N-1}} \, (i = \overline{1,2}).$$

Table 1: Root-Mean-Square Errors for Estimating Accuracy of State ($\sigma_i(200)$)

Coordinate number (i)	LSM	Nonparametric smoothing
1	0.537	0.407
2	0.539	0.479

The corresponding values of the errors of extrapolation of the state vector (see Table 1) were obtained using estimates of unknown vectors of the form (26)–(31).

As can be seen from Table 1, the estimation algorithm with nonparametric smoothing of unknown additive vectors in the object model and in the model of the observation channel allows us to increase the extrapolation accuracy for discrete models with jump parameters.

Conclusions

The solution of the problem of synthesizing stationary and non-stationary robust extrapolators for a linear discrete models with random Markov jump parameters under incomplete information was obtained. The simulation results showed that the application of the robust extrapolation algorithm using the non-parametric smoothing procedure allows one to increase the prediction accuracy.

Acknowledgements

The report study was funded by RFBR according to the research project N 17-08-00920.

References

[1] Bensoussan A. (2018). *Estimation and Control of Dynamical Systems*. Springer International Publishing AG.

[2] Costa E.F., De Saporta B. (2017). Linear Minimum Mean Square Filters for Markov Jump Linear Systems. *IEEE Trans. on Automatic Control*. Vol. **62**. Issue 7, pp. 3567–3572.

[3] Gillijns S., Moor B. (2007). Unbiased Minimum-Variance Input and State Estimation for Linear Discrete-time Systems. *Automatica*. Vol. **43**, pp. 111–116.

[4] Hsien C.S. (2010). On the Optimality of Two-Stage Kalman Filter for Systems with Unknown Input. *Asian Journal of Control*. Vol. **12**, Issue. **4** pp. 510–523.

[5] Janczak D., Grishin Yu. (2006). State Estimation of Linear Dynamic System with Unknown Input and Uncertain Observation Using Dynamic Programming. *Control and Cybernetics*. Vol. **4**, pp. 851–862.

[6] Koshkin G.M., Smagin V.I. (2014). Filtering and Prediction for Discrete Systems with Unknown Input Using Nonparametric Algorithms. *Proc. 10th International Conference on Digital Technologies*. Zilina. Slovakia, pp. 120–124.

[7] Lancaster P. (1969). *Theory of Matrices*. Academic Press, New York.

[8] Li F., Shi P., Wu L. (2016). *Control and Filtering for Semi-Markovian Jump Systems*. Springer, New York.

[9] Liu W. (1985). State Estimation for Discrete-time Markov Jump Linear Systems with Time-correlated Measurement Noise. *Automatica*. Vol. **76**, pp. 266–276.

[10] Lomakina S.S., Smagin V.I. (2005). Robust Filtering for Continuous Systems with Random Jump Parameters and Degenerate Noises in Observation. *Optoelectronics, Instrumentation and Data Processing*. Issue. **2**, pp. 36–43.

[11] Shi P., Boukas E.K., Agarwal R.K. (1999). Kalman Filtering for Continuous-time Uncertain Systems with Markovian Jumping Parameters. *IEEE Transactions on Automatic Control*. Vol. **44**. Issue 8, pp. 1592–1597.

[12] Smagin V.I., Koshkin G.M. (2015). Kalman Filtering and Conrol Algorithms for Systems with Unknown Disturbances and Parameters Using Nonparametric Technique. *Proc. 20th International Conference on Methods and Models in Automation and Robotics (MMAR 2015)*. Miedzyzdroje. Poland, pp. 247–251.

[13] Smagin V.I. (2015). State Estimation for Nonstationary Discrete Systems with Unknown Input Using Compensations. *Russian Physics Journal*. Vol. **58**. Issue **7**, pp. 1010–1017.

[14] Terra M.H., Ishihara J.Y., Cerri J.P. (1985). Robust Estimation for Discrete-time Markovian Jump Linear Systems. *IEEE Transactions on Automatic Control*. Vol. **55**. Issue. **8**, pp. 2065–2071.

[15] Witczak M. (2014). Fault Diagnosis and Fault-Tolerant Control Strategies for Non-Linear Systems. Chapter 2. Unknown Input Observers and Filters. *Lecture Notes in Electrical Engineering. Springer International Publishing*. Switzerland, pp. 19–56.

[16] Yong S.E., Zhu M., Frazzoli E. (2013). Simultaneous Input and State Estimation for Linear Discrete-time Stochastic Systems with Direct Feedthrough. *52nd IEEE Conference on Decision and Control*. Florence. Italy, pp. 7034–7039.

Adaptive prediction of Ornstein-Uhlenbeck process by observations with additive noise

Tatiana V. Dogadova, Vyacheslav A. Vasiliev
National Research Tomsk State University, Tomsk, Russia
e-mail: aurora1900@mail.ru, vas@mail.tsu.ru

Abstract

This paper considers the Ornstein-Uhlenbeck process by observations with additive noise that also satisfies Ornstein-Uhlenbeck equation. The truncated parameter estimation problem of non-observable process with guaranteed accuracy is solved. On the basis of these estimators adaptive predictors of observable process are constructed. Asymptotic property of predictors is established. The presented algorithm works for predictors of any depth.

Keywords: Truncated estimation method, fixed sample size, Ornstein-Uhlenbeck process, guaranteed accuracy, adaptive prediction

Intrduction

One of the important problems of modern applied mathematics is the construction of mathematical models and development of the identification and prediction algorithms with guaranteed accuracy for discrete and continuous time stochastic dynamic systems. Such systems are widely used for the description of databases, for information processing, as well as for mathematical model construction of random processes in economics, financial mathematics, physics, sociology, biology, medicine etc.

The most frequently used for these purposes continuous-time models are the diffusion-type models and the Ito processes. The structure of the abovementioned models implies essential dependence of observations which corresponds to demands for real stochastic processes.

According to Ljung's concept the prediction is a crucial part in constructing complete probabilistic models of dynamical systems (see [1, 2]). A model is considered to be useful if it allows to make predictions of high statistical quality.

Models of dynamical systems often have unknown parameters, which requires estimation in order to build adaptive predictors. The quality of adaptive prediction explicitly depends on the chosen estimators of model parameters. Possible estimation methods include the classic stochastic approximation, maximum likelihood, least squares and sequential estimation methods among others. The first three methods provide estimators with given statistical properties under asymptotic assumptions, when the duration of observations tends to infinity (see, e.g., [3, 4]).

The sequential estimation method makes it possible to obtain estimators with guaranteed accuracy by samples of finite but random and unbounded size (see, e.g., [4]–[11] among others).

Both approaches do not guarantee prescribed estimation accuracy when using samples of non-random finite size and lead up to complicated analytical problems in

adaptive procedures.

However, the more recent truncated sequential estimation method yields estimators with prescribed accuracy by samples of random but bounded size, see [7], [8] among others.

Then the truncated estimation method was introduced in [12]. Truncated estimators were constructed for ratio type multivariate functionals by samples of fixed size and have guaranteed accuracy in the sense of the L_{2m}-norm, $m \geq 1$ (see also [11]). The truncated estimation method is simpler in implementation then the truncated sequential estimation one. At the same time, both methods are very effective in problems of parameter estimation of dynamical systems.

The main aim of the paper is the construction and investigation of adaptive predictors' properties of observable process wich is a sum of two unobservable Ornstein-Uhlenbeck processes. The presented algorithm based on the usage of truncated estimators and works for making predictions of any depth. Similar problems for continuous-time systems were solved in, e.g., [13, 14]. Properties of adaptive optimal control of continuous-time processes constructed on the basis of sequential parameters were considered in [15]. Adaptive optimal predictors for discre-time multivariate system were constructed in [16].

1 Problem statement. Guaranteed parameter estimation of Ornstein-Uhlenbeck process

Consider the estimation problem of the parameter a of the first order stable autoregressive process

$$dx_t = ax_t dt + dw_t, \quad t \geq 0 \tag{1}$$

with the initial value x_0 by oservation of the process y_t with the known parameter λ of the noise θ

$$y_t = x_t + \theta_t, \qquad \theta_t = \lambda\theta_t dt + dv_t, \tag{2}$$

where w_t and v_t are independent standard Wiener processes,θ_0 - initional value for θ, $a < 0, \lambda < 0,\ \lambda^2 \neq a^2$.

Let's substitute an unobservable process x_t in differential equation (1) by the difference $y_t - \theta_t$ and get the eqation

$$dy_t = ay_t dt + dw_t + d\theta_t - a\theta_t dt. \tag{3}$$

Since the parameter λ is known, we have a possibility to exclude the dependent noise θ_t from the equation (3). To this end we integrate it from 0 to t and multiply the result by dt

$$y_t dt = y_0 dt + a\int_0^t y_s ds dt + w_t dt + d\theta_t dt - a\int_0^t \theta_s ds dt.$$

Multiply the obtained equation by λ and subtract it from the equation (3)

$$dy_t - \lambda y_t dt = -\lambda y_0 dt + a\left[y_t - \lambda\int_0^t y_s ds\right] dt + dw_t + \lambda d\theta_t dt$$

$$-\lambda w_t dt - a\left[\theta_t - \lambda\int_0^t \theta_s ds\right] dt.$$

Define $z_t = y_t - \lambda\int_0^t y_s ds$ and then $dz_t = dy_t - \lambda y_t dt$. From the equation $d\theta_t - \lambda\theta_t dt = dv_t$ it follows, that $\theta_t - \lambda\int_0^t \theta_s ds = \theta_0 + v_t$. Last equation can be written in a form

$$dz_t = az_t dt + d(w_t + v_t) - (\lambda w_t + av_t)dt - (\lambda y_0 + a\theta_0)dt.$$

Let us define the difference operator $\delta_h z_t = z_t - z_{t-h}$ with a step $h,\ h > 0$ and apply it to the previous equation

$$d\delta_h z_t = a\delta_h z_t dt + d(\delta_h w_t + \delta_h v_t) - (\lambda\delta_h w_t + a\delta_h v_t)dt. \tag{4}$$

Note that $\delta_h z_t$ is an observable process as well. In view of the fact that $\delta_h z_t$ and model noises are correlated, we construct the correlation (or Yule-Walker) type estimator with the shift h

$$\hat{a}_T = \frac{\int_{2h}^T \delta_h z_{t-h} d\delta_h z_t}{\int_{2h}^T \delta_h z_{t-h}\delta_h z_t dt}. \tag{5}$$

We rewrite the deviation of estimator (5), having replaced $d\delta_h z_t$ by the right hand side of (4)

$$\hat{a}_T - a = \frac{1}{\int_{2h}^T \delta_h z_{t-h}\delta_h z_t dt}\Big[\int_{2h}^T \delta_h z_{t-h} d(\delta_h w_t + \delta_h v_t)$$
$$- \int_{2h}^T \delta_h z_{t-h}(\lambda\delta_h w_t + a\delta_h v_t)dt\Big].$$

Analogously to [17, 18],

$$\frac{1}{T}\left[\int_{2h}^T \delta_h z_{t-h} d(\delta_h w_t + \delta_h v_t) - \int_{2h}^T \delta_h z_{t-h}(\lambda\delta_h w_t + a\delta_h v_t)dt\right] \to 0 \quad \text{a.s.}$$

Taking into account the independence $\delta_h z_{t-h}$ from $\delta_h w_t$ and $\delta_h v_t$,

$$E\left[\int_{2h}^T \delta_h z_{t-h} d(\delta_h w_t + \delta_h v_t) - \int_{2h}^T \delta_h z_{t-h}(\lambda\delta_h w_t + a\delta_h v_t)dt\right]^2$$
$$\le C \cdot E\int_{2h}^T (\delta_h z_{t-h})^2 dt \le C \tag{6}$$

and there exists the limit

$$\sigma_h^2 = \lim_{T\to\infty}\frac{1}{T}\int_{2h}^T \delta_h z_{t-h}\delta_h z_t dt \quad \text{a.s.},$$

where $\sigma_h^2 = \left(1 - \frac{\lambda^2}{a^2}\right)\frac{ea^h - 1}{2a} \neq 0$.

It is easy to verify that

$$\lim_{n\to\infty}\hat{a}_T = a \quad \text{a.s.}$$

and for every $T > 0$ the following inequality holds

$$E\left[\frac{1}{T}\int_{2h}^{T}\delta_h z_{t-h}\delta_h z_t dt - \sigma^2\right]^4 \le \frac{C}{T^2}. \tag{7}$$

The truncated estimator $\tilde{a}_T$ of the parameter a can be defined similar to [12] for some $T_0 > 0$ as

$$\tilde{a}_T = \frac{\int_{2h}^{T}\delta_h z_{t-h} d\delta_h z_t}{\int_{2h}^{T}\delta_h z_{t-h}\delta_h z_t dt}\cdot\chi\left(\left|\int_{2h}^{T}\delta_h z_{t-h}\delta_h z_t dt|\right| \ge T\cdot\log^{-1}T\right). \tag{8}$$

Using (13), (7) and similar to the scheme of the proof for truncated estimators in [12], we get

$$E(\tilde{a}_T - a)^2 \le \frac{C}{T}, \qquad T \ge T_0. \tag{9}$$

By the condition $a < -r$, $r > 0$ the estimator σ_h^2 has the form

$$\sigma_h^2 = \left(1 - \frac{\lambda^2}{\overline{a}^2}\right)\frac{e\overline{a}^h - 1}{2\overline{a}},$$

where $\overline{a} = \mathrm{proj}_{(-\infty,-r]}\tilde{a}_s$ and satisfy the condition

$$E(\sigma_h^2 - \sigma^2)^2 \le \frac{C}{T}, \qquad T \ge T_0.$$

Without a priory information about a, the truncated estimation method can be applied for estimation σ_h^2 .

2 Adaptive prediction

Consider the model (1), (2). The purpose is to construct an adaptive predictor for y_t by observations $y^{t-u} = (y_s)_{0\le s\le t-u}$. Here $u > 0$ - is a fixed time delay.

Using the solution of the equation (1), we get

$$x_t = \mu x_{t-u} + \xi_{t,t-u}, \quad t \ge u, \tag{10}$$

where $\xi_{t,t-u} = \int\limits_{t-u}^{t} e^{a(t-s)}dw_s$, $\mu = e^{au}$.

Define

$$\mu_s = e^{\hat{a}_s u}, \quad s \ge 0. \tag{11}$$

Here

$$\hat{a}_s = \mathrm{proj}_{(-\infty,0]}\tilde{a}_s,$$

$\hat{a}_s$ is a projection of the truncated estimator $\tilde{a}_s$ of the parameter a, defined in (7).

It can be shown that

$$E(\mu_t - \mu)^{2p} \le \frac{C}{t^p}, \quad p \ge 1. \tag{12}$$

Replacing x_t in the formula (9) using (2) we get

$$y_t = \mu y_{t-u} + \xi_{t-u,t} + \theta_t - \mu\theta_{t-u},$$

Introduce the notation

$$\eta_{t-u,t} = \int_{t-u}^{t} e^{\lambda(t-s)} dw_s, \quad \overline{\xi}_{t-u,t} - e^{\lambda u}\xi_{t-2u,t-u}, \quad \overline{\eta}_{t-u,t} - e^{\lambda u}\xi_{t-2u,t-u}.$$

and

$$z_t = y_t - e^{\lambda u} y_{t-u}.$$

The function z_t satisfies the equation

$$z_t = \mu z_{t-u} + \overline{\xi}_{t-u,t} + \overline{\eta}_{t-u,t} - \mu\overline{\eta}_{t-2u,t-u}. \tag{13}$$

Applying operator of conditional mathematical expectation $E(\cdot|y^{t-3u})$ to the last equation we get

$$E(z_t|y^{t-3u}) = \mu E(z_{t-u}|y^{t-3u}).$$

By the definition of z_t we have

$$E(z_t|y^{t-3u}) = E(y_t|y^{t-3u}) - e^{\lambda u} E(y_{t-u}|y^{t-3u}).$$

Let us define $s_i(t) = E(y_t|y^{t-iu})$, $i = \overline{1,3}$.
The equation for optimal predictions $s_i(t)$, $i = \overline{1,3}$, has the form

$$s_3(t) = (e^{au} + e^{\lambda u})s_2(t) + e^{(a+\lambda)u} s_1(t).$$

Define adaptive predictors $\hat{s}_i(t)$, $i = \overline{1,3}$. The equation for $\hat{s}_i(t)$, is constructed with truncated estimators instead of unknown parameters

$$\hat{s}_3(t) = (e^{\hat{a}_{t-3u}u} + e^{\lambda u})\hat{s}_2(t) + e^{(\hat{a}_{t-3u}+\lambda)u}\hat{s}_1(t).$$

Prediction errors can be written as

$$e_i(t) = s_i(t) - \hat{s}_i(t), \qquad i = \overline{1,3}.$$

It can be shown that

$$\overline{\lim}_{t\to\infty} Ee_i^2(t) < \infty, \qquad i = \overline{1,3}.$$

In the conclusion we note that obtained property for this model probably can not be improved in view of complicated structure of noise dependence. At the same time this property reflects proximity of adaptive and optimal predictors in L_2 - metric, which is important in analytical investigations and practical applications.

Acknowledgment

The work is supported by the Russian Science Foundation under grant No. 17-11-01049.

References

[1] Ljung L. (1987). *System identification theory for user.* Englewood Cliffs : Prentice-Hall.

[2] Ljung L., Söderström T. (1983). *Theory and practice of recursive identification.* Massachusets : The MIT Press.

[3] T. W. Anderson, *The Statistical Analysis of Time Series.* New York, NY: John Wiley and Sons, Inc., 1971.

[4] H. Robbins and S. Monro (1951). A Stochastic Approximation Method. *Ann. of Math. Stat.* Vol. **22**, pp. 400–407.

[5] L. Galtchouk and V. Konev (2001). On sequential estimation of parameters in semimartingale regression models with continuous time parameter. *Ann. of Stat.* Vol. **29**, pp. 1508–1536.

[6] U.Küchler and V. A. Vasiliev (2001). On Sequential Parameter Estimation for Some Linear Stochastic Differential Equations with Time Delay. *Sequential Anal.* Vol. **20**, pp. 117–146.

[7] V. V. Konev and S. M. Pergamenshchikov (1990). Truncated Sequential Estimation of the Parameters in Random Regression. *Sequential Anal.* Vol. **9**, pp. 19–41.

[8] V. V. Konev and S. M. Pergamenshchikov (1990). On Truncated Sequential Estimation of the Drifting Parameter Mean in the First Order Autoregressive Model. *Sequential Anal.* Vol. **9**, pp. 193–216.

[9] U. Küchler and V. A. Vasiliev (2010). On Guaranteed Parameter Estimation of a Multiparameter Linear Regression Process. *Automatica.* Vol. **46**, pp. 637–646.

[10] R. S. Liptser and A. N. Shiryaev, *Statistics of Random Processes,* Vol. 1. New York, NY: Springer, 1977.

[11] V. A. Vasiliev (2014). Guaranteed Estimation of Logarithmic Density Derivative By Dependent Observations. In: *Topics in Nonparametric Statistics*, M. G. Akritas *et al., Ed., Springer Proceedings in Mathematics and Statistics*. New York, NY: Springer. Vol. **74**, pp. 351–360.

[12] Vasiliev V. A. (2014). A truncated estimation method with guaranteed accuracy. *Ann. Inst. Stat. Math.* Vol. **66**, no. 1. pp. 141-163.

[13] T. V. Dogadova and V. A. Vasiliev (2017). Adaptive prediction of stochastic differential equations with unknown parameters. *Vestnik of Tomsk State University, J. Control Comput. Sci.*. Vol. **38**, pp. 17–23.

[14] T. V. Dogadova and V. A. Vasiliev (2017). On Adaptive Optimal Prediction of Ornstein–Uhlenbeck Process. *Appl. Math. Sci.* Vol. **11**, pp. 591–600.

[15] Küchler U., Vasiliev V. (2013). On a certainty equivalence design of continuous-time stochastic systems. *SIAM Journal of Control and Optimization*. Vol. **51**, no. 2. pp. 938-964.

[16] M. I. Kusainov and V. A. Vasiliev (2015). On optimal adaptive prediction of multivariate autoregression. *Sequential Anal.* Vol. **34**, pp. 211-234.

[17] U. Küchler, Vasiliev V. (2007). On parameter estimation of stochastic delay differential equations with guaranteed accuracy by noisy observations. *Journal of Statistical Planning and Inference. Elsevier Science*, 137, 3007-3023, doi: 10/1016j.jspi.2006.12.001.

[18] Vasiliev V. A., Konev V. V. (1987). On sequential identification of linear dynamic systems in continuous time by noisy in observations. *Problems of Control and Information Theory*. Vol. **16**, no. 2. pp. 101–112.

Parameter estimation with guaranteed accuracy for AR(1) by noised observations

Yulia B. Burkatovskaya[1,2] and Vyacheslav A. Vasiliev[1]
[1] *Tomsk State University, Tomsk, Russian Federation*
[2] *Tomsk Polytechnic University, Tomsk, Russian Federation*
e-mail: tracey@tpu.ru, vas@mail.tsu.ru

Abstract

This paper presents a truncated estimator of the dynamic parameter of a stable AR(1) process by observations with additive noise. The estimator is constructed by a sample of a fixed size and it has a known upper bound of the mean square deviation. Cases of known and unknown variance of observation noise are considered.

Keywords: autoregressive process, fixed sample size, guaranteed accuracy, observations with noise.

1 Introduction and problem statement

Development of parameter estimation methods of dynamic systems by samples of finite or fixed size is very important in statistical problems such that model construction and various adaptive problems (prediction, control, filtration etc.).

One of the possibilities for finding estimators with the guaranteed quality of inference using a sample of fixed size is provided by the approach of truncated estimation. Truncated estimators were constructed in [9] for ratio type multivariate functionals by a fixed-size sample. They have guaranteed accuracy in the sense of the L_{2m}-norm, $m \geq 1$. This fact allows one to obtain desired non-asymptotic and asymptotic properties of the estimators. The truncated estimation method was developed in [1] and others for parameter estimation problems in discrete-time dynamic models. Solutions of some non-asymptotic parametric and non-parametric problems can be found also in [4], [8], [5], [6], among others. In particular, [8] established the minimax optimality of the least-squares estimator of the dynamic parameter in AR(1) model.

In this paper, the truncated estimation method introduced in [9] is applied for the parameter estimation of AR(1) by additively-noised observations with unknown noise variance (another applications of this method can be found, e.g., in [2], [3]).

Consider the estimation problem of the parameter λ of the scalar first-order autoregressive process $(x_n)_{n\geq 0}$ satisfying the equation

$$x_n = \lambda x_{n-1} + \xi_n, \quad n \geq 1 \tag{1}$$

by observations

$$y_n = x_n + \eta_n, \quad n \geq 0. \tag{2}$$

Process (1) is supposed to be stable, i.e. $|\lambda| < 1$. Introduce the notation $\zeta = (x_0, \xi_1, \eta_0.)$ The processes (ξ_n), (η_n) and x_0 are supposed to be mutually independent;

noises ξ_n and η_n form sequences of i.i.d. random variables such that $E\zeta = 0, E||\zeta||^4 < \infty$. Denote $\sigma^2 = E\eta_0^2$. We assume that the variance of ξ_1 is known. Then without loss of generality we put $E\xi_1^2 = 1$.

The main aim of the paper is to construct truncated estimators of $\lambda \in (-1, 1)$ with guaranteed accuracy in the mean square sense by sample of fixed size. Cases of both known and unknown values of σ^2 will be considered.

A similar problem has been solved in, e.g., [10] on the basis of the sequential approach (when the sample size is a random value determined by a special stopping rule) for $\lambda \in (-1, 0) \cap (0, 1)$

2 Parameter estimation of AR(1) with known noise variance

To estimate the parameter λ, we use the correlation method. To this end, we obtain from the system (1), (2) the recurrent equation for the observed process $y = (y_n)_{n\geq 0}$:

$$\begin{array}{c} y_n = \lambda y_{n-1} + \delta_n, \quad n \geq 1, \\ \delta_n = \xi_n + \eta_n - \lambda\eta_{n-1}. \end{array} \tag{3}$$

Due to the dependence of noises δ_n, the least squares estimator (LSE) of λ obtained from equation (3) is asymptotically biased, see, e.g., [7], [10]. Equation (3) implies the following formula for correlations of the process (y_n):

$$E_\lambda y_n y_{n-1} = \lambda E_\lambda (y_{n-1}^2 - \sigma^2), \quad n \geq 1.$$

Hence, the consistent correlation estimator $\hat{\lambda}_n$ of λ has the following form (see [7])

$$\hat{\lambda}_{n,\sigma} = \frac{\sum\limits_{k=1}^{n} y_k y_{k-1}}{\sum\limits_{k=1}^{n} (y_{k-1}^2 - \sigma^2)}, \quad n \geq 1. \tag{4}$$

It is easy to verify that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (y_{k-1}^2 - \sigma^2) = \frac{1}{1-\lambda^2} > 1 \quad P_\lambda - a.s. \tag{5}$$

Thus, according to the general procedure described in [9], it is reasonable to construct the truncated estimator $\tilde{\lambda}_n$ of λ as follows:

$$\tilde{\lambda}_n = \hat{\lambda}_n \cdot \chi(\sum_{k=1}^{n} (y_{k-1}^2 - \sigma^2) \geq hn), \quad n \geq 1, \tag{6}$$

where $h \in (0, 1)$ and $\chi(A)$ is the indicator of the set A.

The following theorem gives the first main result of this paper.

Theorem 1. *Assume model (1), (2). Then for every* $|\lambda| < 1$ *and* $n \geq 1$, *estimator (7) has the property*

$$E_\lambda(\tilde{\lambda}_n - \lambda)^2 \leq \frac{C}{n}. \tag{7}$$

The proofs of theorems and lemmas are given in Section 5.

3 Parameter estimation of AR(1) with unknown noise variance

To estimate $\lambda \in (-1, 1)$, we use an adaptive modification of estimator (5):

$$\lambda_n^* = \frac{\frac{1}{n}\sum\limits_{k=1}^{n} y_k y_{k-1}}{\frac{1}{n}\sum\limits_{k=1}^{n} y_{k-1}^2 - \sigma_n^2}, \quad n > 1. \tag{8}$$

Taking into account (6), we construct the estimator σ_n^2 of σ^2 as follows

$$\sigma_n^2 = \frac{1}{n}\sum_{k=1}^{n} y_{k-1}^2 - \frac{1}{1-\lambda_n^2}, \quad n > 1 \tag{9}$$

where λ_n is the pilot estimator of λ

$$\lambda_n = \mathrm{proj}_{[-1,1]}\breve{\lambda}_n, \quad n > 1, \tag{10}$$

$$\breve{\lambda}_n = \frac{\sum\limits_{k=2}^{n} y_k y_{k-2}}{\sum\limits_{k=2}^{n} y_{k-1} y_{k-2}} \cdot \chi(|\sum_{k=2}^{n} y_{k-1} y_{k-2}| \geq H_n), \quad n > 1. \tag{11}$$

Here we put $H_n = n(\log n)^{-1}$. According to the general truncated estimation method [9], the multiplier $(\log n)^{-1}$ in the definition of H_n can be any other slowly-decreasing function.

It should be noted that the estimator (10) is constructed on the bases of the correlation (Yule-Walker type) estimator which can not be used if $\lambda = 0$ (see Lemma 1 below). Our main aim is to construct an estimator of λ without this restriction.

Taking into account (10), estimator (9) can be written in the form

$$\lambda_n^* = (1 - \lambda_n^2)\frac{1}{n}\sum_{k=1}^{n} y_k y_{k-1}, \quad n > 1. \tag{12}$$

Lemma 1. *Assume that in model (1), (2),* $E||\zeta||^8 < \infty$. *Then estimator (10) for every* $\lambda \in (-1, 0) \cup (0, 1)$ *and* $n > 1$ *has the following property*

$$E_\lambda(\lambda_n - \lambda)^2 \leq \frac{C_1}{n} + C_2 \frac{\log^4 n}{n^2}.$$

This lemma makes possible to obtain the main result of the section.

Theorem 2. *Assume that in model (1), (2),* $E||\zeta||^8 < \infty$. *Then for every* $|\lambda| < 1$ *and* $n > 1$, *estimator (12) satisfies the following condition*

$$E_\lambda(\lambda_n^* - \lambda)^2 \leq \frac{C}{n} + C \frac{\log^4 n}{n^2}.$$

4 Simulation Results and Discussion

We conducted numerical simulation of the proposed estimation algorithm. For every set of the parameters, the experiment was performed 100 times, the number of observations is equal to 100, the parameter of the procedure $h = 0,5$. Table 1 presents the results of simulation. Here λ and σ are the parameters of model (1), $\tilde{\lambda}_n$ and λ_n^* are the mean estimators of the parameter λ when the noise variance σ^2 is supposed to be known and unknown, correspondingly; $\tilde{d}_n$ and d_n^* are sample standard errors of the corresponding estimators.

One can see that $\tilde{d}_n < d_n^*$ in all experiments; thus, if the noise variance is unknown then the standard error increases at least twice (if $\lambda = 0,5$); but d_n^* can be fully ten times larger than $\tilde{d}_n$, if $\lambda = 0,9$. Both deviations increase with the grow of σ^2, as one should expect; besides, $\tilde{d}_n$ decreases and d_n^* increases with the increase of λ.

5 Proofs

5.1 *Proof of Theorem 1*

To investigate the non-asymptotic properties of $\tilde{\lambda}_n$ we use the following representation of the deviation

$$\tilde{\lambda}_n - \lambda = \frac{f_n}{g_n} \cdot \chi(|g_n| \geq h) - \lambda \chi(|g_n| < h), \tag{13}$$

where

$$f_n = \frac{1}{n} \sum_{k=1}^{n} [y_{k-1}(\xi_k + \eta_k) - \lambda(y_{k-1}\eta_{k-1} - \sigma^2)],$$

$$g_n = \frac{1}{n} \sum_{k=1}^{n} (y_{k-1}^2 - \sigma^2).$$

It can be directly verified that for $|\lambda| < 1$

$$E_\lambda f_n^2 \leq \frac{I^{-1}(\lambda, \overline{\sigma})}{n}. \tag{14}$$

Table 1: Simulation results

λ	σ^2	$\tilde{\lambda}_n$	$\tilde{d}_n$	λ_n^*	d_n^*
0,5	0,09	0,477	0,0092	0,452	0,0294
0,5	0,25	0,492	0,0111	0,490	0,0314
0,5	0,49	0,487	0,0150	0,470	0,0488
0,5	0,81	0,475	0,0229	0,418	0,0795
0,5	1	0,473	0,0419	0,424	0,0953
0,8	0,09	0,786	0,0046	0,796	0,0465
0,8	0,25	0,794	0,0054	0,854	0,0793
0,8	0,49	0,786	0,0054	0,789	0,0737
0,8	0,81	0,772	0,0120	0,765	0,1435
0,8	1	0,788	0,0122	0,797	0,1590
0,9	0,09	0,876	0,0038	0,865	0,0772
0,9	0,25	0,889	0,0018	0,913	0,0596
0,9	0,49	0,888	0,0030	0,910	0,1044
0,9	0,81	0,874	0,0044	0,886	0,1822
0,9	1	0,891	0,0028	0,891	0,1780

Introduce the notation $g = 1/(1-\lambda^2)$. Then, using a representation

$$g_n - g = \frac{1}{n}\sum_{k=1}^{n}(x_{k-1}^2 - \sigma^2) + \frac{2}{n}\sum_{k=1}^{n} x_{k-1}\eta_{k-1} + \frac{1}{n}\sum_{k=1}^{n}(\eta_{k-1}^2 - \sigma^2)$$

and the following formula (see, e.g., the proof of Theorem 2 in [9])

$$\frac{1}{n}\sum_{k=1}^{n}(x_{k-1}^2 - g) = \frac{g}{n}\cdot[x_0^2 - x_n^2 + 2\lambda\sum_{k=1}^{n} x_{k-1}\xi_k + \sum_{k=1}^{n}(\xi_k^2 - 1)],$$

it is easy to prove that

$$E_\lambda(g_n - g)^2 \le \frac{C_0}{n}, \qquad n \ge 1. \tag{15}$$

Further, similar to [9] using the Chebyshev inequality we estimate

$$P_\lambda(|g_n| < h) \le P_\lambda(|g_n - g| > g - h) \le \frac{E_\lambda(g_n - g)^2}{(g-h)^2} \le \frac{C_0}{(1-h)^2 n}, \qquad n \ge 1. \tag{16}$$

Using (13–16), we estimate

$$E_\lambda(\tilde{\lambda} - \lambda)^2 \le \frac{1}{h^2}E_\lambda f_n^2 + P_\lambda(|g_n| < h) \le \frac{I^{-1}(\lambda, \overline{\sigma})}{h^2 n} + \frac{E_\lambda(g_n - g)^2}{(g-h)^2} \le \frac{C}{n}$$

and obtain assertion (7).

5.2 *Proof of Lemma 1*

The proof of Lemma 1 is similar to the proof of the second assertion of Theorem 1 in [9].

Definition (10) of λ_n implies

$$E_\lambda(\lambda_n - \lambda)^2 \le E_\lambda(\breve{\lambda}_n - \lambda)^2.$$

Introduce the following notations

$$f_n = \frac{1}{n}\sum_{k=2}^{n} y_{k-2}\delta_k, \quad g_n = \frac{1}{n}\sum_{k=1}^{n} y_{k-1}y_{k-2}, \quad g = \frac{\lambda}{1-\lambda^2}, \quad h_n = (\log n)^{-1}.$$

By the definition of $\breve{\lambda}$ in (11), its deviation has the form

$$\breve{\lambda}_n - \lambda = \frac{f_n}{g_n}\cdot\chi(|g_n| \ge h_n) - \lambda\cdot\chi(|g_n| < h_n) = \frac{f_n}{g}\cdot\chi(|g_n| \ge h_n)$$
$$+\frac{f_n(g-g_n)}{gg_n}\cdot\chi(|g_n| \ge h_n) - \lambda\cdot\chi(|g_n| < h_n) = J_1 + J_2 + J_3.$$

Using the Cauchy-Schwarz-Bunyakovsky and Chebyshev's inequalities, estimate the second moments of these summands:

$$E_\lambda J_1^2 \le CE_\lambda f_n^2, \quad E_\lambda J_2^2 \le \frac{1}{g^2h_n^2}\sqrt{E_\lambda f_n^4 E_\lambda(g_n-g)^4}, \quad E_\lambda J_3^2 \le h_n^{-4}E_\lambda(g_n-g)^4.$$

In view of the structure of the function f_n it is easy to verify that $E_\lambda f_n^4 \le C/n^2$.

By the definition of g_n we have

$$g_n - g = \frac{1}{n}\sum_{k=1}^{n} y_{k-1}y_{k-2} - \frac{\lambda}{1-\lambda^2} = \frac{1}{n}\sum_{k=1}^{n} x_{k-1}x_{k-2} - \frac{\lambda}{1-\lambda^2}$$
$$+\frac{1}{n}\sum_{k=1}^{n} x_{k-1}\eta_{k-2} + \frac{1}{n}\sum_{k=1}^{n}\eta_{k-1}x_{k-2} + \frac{1}{n}\sum_{k=1}^{n}\eta_{k-1}\eta_{k-2} = \lambda\left(\frac{1}{n}\sum_{k=1}^{n} x_{k-2}^2 - \frac{1}{1-\lambda^2}\right)$$
$$+\frac{1}{n}\sum_{k=1}^{n} x_{k-1}\eta_{k-2} + \frac{1}{n}\sum_{k=1}^{n}(\eta_{k-1}+\xi_{k-1})x_{k-2} + \frac{1}{n}\sum_{k=1}^{n}\eta_{k-1}\eta_{k-2}.$$

Using this representation, it is easy to verify similarly the proof of Theorem 1 that

$$E_\lambda(g_n - g)^4 \le \frac{C}{n^2}.$$

Thus we have

$$E_\lambda J_1^2 \le C\frac{1}{n}, \quad E_\lambda J_2^2 \le C\frac{\log^2 n}{n^2}, \quad E_\lambda J_3^2 \le C\frac{\log^4 n}{n^2}.$$

5.3 *Proof of Theorem 2*

Introduce the following notations

$$\Delta_n = (\lambda^2 - \lambda_n^2)\frac{\lambda}{1-\lambda^2} + (1-\lambda_n^2)\lambda\frac{1}{1-\lambda^2}\left\{\frac{1}{n}(x_0^2 - \lambda^2 x_{k-1}^2) + \frac{2\lambda}{n}\sum_{k=2}^{n} x_{k-2}\xi_{k-1} + \frac{1}{n}\sum_{k=2}^{n}(\xi_{k-1}^2 - 1) - \frac{1}{n}\right\} + (1-\lambda_n^2)\frac{1}{n}\sum_{k=1}^{n}[\lambda x_{k-1}\eta_{k-1} + y_{k-1}(\xi_k + \eta_k)].$$

Definition (12) of the estimator λ_n^* and equation (3) imply

$$\lambda_n^* = (1-\lambda_n^2)\frac{1}{n}\sum_{k=1}^{n}[\lambda y_{k-1}^2 + y_{k-1}(\xi_k + \eta_k) - \lambda y_{k-1}\eta_{k-1}]$$

$$= (1-\lambda_n^2)\left\{\lambda\frac{1}{n}\sum_{k=1}^{n} x_{k-1}^2 + \frac{1}{n}\sum_{k=1}^{n}[\lambda x_{k-1}\eta_{k-1} + y_{k-1}(\xi_k + \eta_k)]\right\}$$

$$= (1-\lambda_n^2)\frac{\lambda}{1-\lambda^2} + (1-\lambda_n^2)\lambda\left[\frac{1}{n}\sum_{k=1}^{n} x_{k-1}^2 - \frac{1}{1-\lambda^2}\right]$$

$$+(1-\lambda_n^2)\frac{1}{n}\sum_{k=1}^{n}[\lambda x_{k-1}\eta_{k-1} + y_{k-1}(\xi_k + \eta_k)] = \lambda + (\lambda^2 - \lambda_n^2)\frac{\lambda}{1-\lambda^2}$$

$$+(1-\lambda_n^2)\lambda\frac{1}{1-\lambda^2}\left\{\frac{1}{n}(x_0^2 - \lambda^2 x_{k-1}^2) + \frac{2\lambda}{n}\sum_{k=2}^{n} x_{k-2}\xi_{k-1} + \frac{1}{n}\sum_{k=2}^{n}(\xi_{k-1}^2 - 1) - \frac{1}{n}\right\}$$

$$+(1-\lambda_n^2)\frac{1}{n}\sum_{k=1}^{n}[\lambda x_{k-1}\eta_{k-1} + y_{k-1}(\xi_k + \eta_k)] = \lambda + \Delta_n.$$

Thus the mean square deviation of the estimator λ_n^* has the following form

$$E_\lambda(\lambda_n^* - \lambda)^2 = E_\lambda\Delta_n^2 \cdot \chi(\lambda = 0) + E_\lambda\Delta_n^2 \cdot \chi(\lambda \neq 0) =: I_1 + I_2,$$

where

$$I_1 = E_\lambda((1-\lambda_n^2)\frac{1}{n}\sum_{k=1}^{n}[y_{k-1}(\xi_k + \eta_k)])^2 \cdot \chi(\lambda = 0)$$

$$= E_\lambda((1-\lambda_n^2)\frac{1}{n}\sum_{k=1}^{n}(\xi_{k-1} + \eta_{k-1})(\xi_k + \eta_k))^2 \cdot \chi(\lambda = 0),$$

$$I_2 = E_\lambda\Delta_n^2 \cdot \chi(\lambda \neq 0).$$

From assumptions of Theorem 2 it follows $I_1 \leq C/n$. In view of Lemma 1 and the property $|\lambda_n + \lambda| \leq 2$, we have

$$I_2 \leq E_\lambda\left(\frac{2|\lambda_n - \lambda|}{1-\lambda^2}\chi(\lambda \neq 0) + \frac{1}{1-\lambda^2}\left\{\frac{1}{n}(x_0^2 + x_{k-1}^2) + \frac{2}{n}\left|\sum_{k=2}^{n} x_{k-2}\xi_{k-1}\right| + \frac{1}{n}\left|\sum_{k=2}^{n}(\xi_{k-1}^2 - 1)\right| + \frac{1}{n}\right\} + \frac{1}{n}\left|\sum_{k=1}^{n} x_{k-1}\eta_{k-1}\right| + \frac{1}{n}\left|\sum_{k=1}^{n} y_{k-1}(\xi_k + \eta_k)\right|\right)^2$$

$$\leq CE_\lambda(\lambda_n - \lambda)^2\chi(\lambda \neq 0) + \frac{C}{n} + \frac{C}{n^2} \leq \frac{C}{n} + C\frac{\log^4 n}{n^2}.$$

Acknowledgment

This study is supported by the RSF (project 17-11-01049).

References

[1] Dogadova, T. V., and V. A. Vasiliev. 2014. Guaranteed parameter estimation of stochastic linear regression by sample of fixed size. *Tomsk State University Journal of Control and Computer Science*. Vol. **26(1)**, pp. 39–52.

[2] Tatiana V. Dogadova, Marat I. Kusainov and Vyacheslav V. Vasiliev. 2017. *Truncated estimation method and applications. Serdica. Mathematical Journal Bulgarian Academy of Sciences Institute of Mathematics and Informatics*. Vol. **43**, pp. 221–266.

[3] Kusainov, M. I., and V. A. Vasiliev. 2015. On optimal adaptive prediction of multivariate autoregression. *Sequential Analysis: Design Methods and Applications*. Vol. **35(2)**, pp. 211–234.

[4] Mikulski, P. W., and M. J. Monsour. 1991. Optimality of the maximum likelihood estimator in first-order autoregressive processes. *Journal of Time Series Analysis*. Vol. **12(3)**, pp. 237–253.

[5] Roll, J., A. Nazin, and L. Ljung. 2002. A Non-Asymptotic Approach to Local Modelling. *The 41st IEEE CDC, Las Vegas, Nevada, USA, 10-13 Dec. 2002 (Regular paper)*. Predocumentation is available at http://www.control.isy.liu.se/research/reports/2003/2482.pdf

[6] Roll J, Nazin A, Ljung L. 2005. Non-linear System Identification Via Direct Weight Optimization. *Automatica, Journal of IFAC. Special Issue on "Data-Based Modelling and System Identification"*. Vol. **41(3)**, pp. 475–490. Predocumentation is available at http://www.control.isy.liu.se/research/reports/2005/2696.pdf

[7] Shneeweiss, H. 1976. Consistent estimation of a regression with error in the variables. *Metrika*. Physica-Verlag, Wien. Vol. **23**, pp. 101–115.

[8] Shiryaev, A. N., and V. G. Spokoiny. 2000. *Statistical Experiments and Decisions. Asymptotic Theory*. Singapore, New Jersey, London, Hong Kong: World Scientific.

[9] Vasiliev, V. 2014. A Truncated Estimation Method with Guaranteed Accuracy. *Annals of Institute of Statistical Mathematics*. Vol. **66**, pp. 141–163.

[10] Vasiliev V. A., and V. V. Konev 1982. The Sequential Parameters Estimation of Dynamic Systems with Noisy Observations. *Izvestia Acad. USSR, Technical Cybernetics*. Vol. **6**, pp. 145–154.

Minimum gamma-divergence estimation for non-homogeneous data with application to ordered probit model

Daniil V. Lisitsin and Andrey G. Usol'tsev
Novosibirsk State Technical University, Novosibirsk, Russia
e-mail: lisitsin@ami.nstu.ru,usol1996@gmail.com

Abstract

In paper, optimality properties of the minimum gamma-divergence estimator (MGDE) relative to weighted L_2-norm of Hampel's influence function, and also Shurygin's model of the point Bayesian contamination under theory of parameter estimation from multivariate non-homogeneous data are considered. Influence functions of the MGDE and close to it of the minimum beta-divergence estimator and the generalized radical estimators for the ordered probit model are compared. The MGDE applicability is evaluated using a simulated dataset.

Keywords: *M*-estimator, robust estimation, influence function, gamma-divergence, beta-divergence, non-homogeneous data, ordered probit model.

Introduction

The classical statistic procedures are based on a number of assumptions which cannot be fulfilled in practice. Under such conditions a lot of widespread statistic procedures lose their positive qualities. For instance, the procedures, which rest on the maximum likelihood estimator (MLE). However, this problem can be solved by using robust estimators. Generally, robustness theory has been developed for the quantitative continuous random variables modeling [3, 12]. Much less attention is paid to the modeling of ordinal regression, and existing approaches are often based on semi-heuristic nature. Few papers are devoted to robust parameter estimation of the cumulative link model and the ordinal probit model as its particular case (see, for example [4, 10, 19]).

Previously, we developed the general theory of asymptotically optimal estimation of unknown model parameters from multivariate non-homogeneous incomplete data (see [5]). At the bottom of this theory we find synthesis of approach by F.R. Hampel [8] which is associated with the influence function and approach by A.M. Shurygin [18] which is associated with the point Bayesian contamination model. Resulting methods are robust against the model misspecification, estimators often have redescending property [17, 18]. This theory is applied to cases with non-homogeneous quantitative (including count), qualitative, mixed data, and also in the presence of missing data (see references in [3]). The application to qualitative response regression models is considered in [12].

The very robust minimum gamma-divergence (or logarithmic density power divergence) estimator (MGDE) is used for a parameter estimation of distributions [11],

regression [6], heteroscedactic regression [14], dichotomous logistic regression [9]. In the latter two cases the data are non-homogeneous. Many well-known robust estimators of location are special versions of the MGDE for suitable model (see, for example, [16, 17]). These are the estimators associated to names of Tukey (biweight), Andrews (sine), Huber (skipped mean), Welsch, Smith, Bernoulli, Charbonnier (generalized version). For example, the latter case is related to the Student's t distribution.

In connection with the above, it is important to find out optimality properties of the MGDE, corresponding results are given in Section 1. In Section 2, we apply the MGDE to the ordered probit model, namely, we compare influence functions of the MGDE and similar estimators, also present the result for one simulated dataset.

1 Optimality properties of the MGDE

Let n-dimensional independent random variables $\zeta_i = (\zeta_{i1}, \ldots, \zeta_{in})^{\mathsf{T}}$, $i = 1, \ldots, N$, have an assumed (or ideal) probability density functions (p.d.f.'s) $g_i(z_i|\phi)$, $z_i \in R^n$, with respect to a σ-finite measure μ, ϕ is p-vector of parameters.

M-estimate $\hat{\phi}$ of vector ϕ is obtained from the observations $\tilde{\zeta}_i$, $i = 1, \ldots, N$, of random variables ζ_i, $i = 1, \ldots, N$, by means of a solution of the system of estimating equations

$$\sum_{i=1}^{N} \psi_i(\tilde{\zeta}_i, \hat{\phi}) = 0,$$

where $\psi_i(\tilde{\zeta}_i, \hat{\phi})$ is p-dimensional estimating function satisfying further condition

$$\mathbf{E}\, \psi_i(z_i, \phi) = 0, \quad i = 1, \ldots, N, \tag{1}$$

$\mathbf{E}$ is expectation under the assumed p.d.f. [1]. Alternative (but no equivalent) way of definning M-estimate is optimization, namely

$$\hat{\phi} = \arg \min_{\tilde{\phi}} \sum_{i=1}^{N} \rho_i(\tilde{\zeta}_i, \tilde{\phi}),$$

where $\rho_i(\tilde{\zeta}_i, \tilde{\phi})$ is a loss function [1]. We obtain the first way by choosing $\psi_i(z_i, \tilde{\phi})$ proportional to the gradient of $\rho_i(z_i, \tilde{\phi})$.

Robust estimates have high quality not only in the assumed distribution, but in the case of deviations from it. One form of deviation is defined by the contaminated distribution such that the real p.d.f. of observations is determined by the mixture model $(1-\varepsilon)g_i(z_i|\phi)+\varepsilon h_i(z_i)$, where $0 \le \varepsilon < 1$ is an amount of contamination, $h_i(z_i)$ is a contamination p.d.f.

One of the major indicators of estimator's robustness is the influence function [8]. In our case, for M-estimator under certain regularity conditions, the influence function for the ith observation take the form [5]

$$\mathrm{IF}_i(z_i, \psi) = \mathrm{M}^{-1}\psi_i(z_i, \phi),$$

where $\psi = (\psi_1^\mathsf{T}, \ldots, \psi_N^\mathsf{T})^\mathsf{T}$,

$$\mathrm{M} = -\sum_{i=1}^{N} \frac{\partial}{\partial \tilde{\phi}^\mathsf{T}} \mathbf{E}\, \psi_i(z_i, \tilde{\phi})\Big|_{\tilde{\phi}=\phi} = \sum_{i=1}^{N} \int\limits_{R^n} \psi_i(z_i, \phi) \frac{\partial g_i(z_i|\phi)}{\partial \phi^\mathsf{T}}\, d\mu$$

is non-singular $p \times p$ matrix.

Indicator of estimation badness can be written as square of the weighted L_2-norm of the influence function [5]

$$\Lambda_s(\psi) = \sum_{i=1}^{N} \int\limits_{R^n} \mathrm{IF}_i^\mathsf{T}(z_i, \psi)\, W\, \mathrm{IF}_i(z_i, \psi) s_i(z_i|\phi)\, d\mu, \tag{2}$$

where $s = (s_1, \ldots, s_N)^\mathsf{T}$, $s_i(z_i|\phi) > 0$ is weight function, $W = W(\phi)$ is a symmetric positive definite weight matrix of size $p \times p$ (under some conditions W can provide invariance of Λ_s to one-to-one differentiable parameter transformation).

Optimal estimating function is a solution of minimization problem [5]:

$$\psi_s^* = \arg\min_{\psi} \Lambda_s(\psi)$$

under constraints (1) and has the form

$$\psi_{s,i}^*(z_i, \phi) = C \left[\frac{\partial}{\partial \phi} \ln g_i(z_i|\phi) + b_i \right] \frac{g_i(z_i|\phi)}{s_i(z_i|\phi)}, \tag{3}$$

where $C = C(\phi)$ is insignificant non-singular matrix, $b_i = b_i(\phi)$ is determined from condition (1).

Also, indicator (2) can be interpreted in accordance with the model of the point Bayesian contamination [18].

Consider a version non-homogeneous point Bayesian contamination model (see [14]). There is a series of samples with random point contamination. A random contamination point in the ith observation has a p.d.f. $\chi_i(z_i|\phi)$, $z_i \in R^n$, with respect to measure μ. The ith observation is the only one contaminated in all samples and has a random contamination point z_{ki}^* in the kth sample and an amount of contamination equal to ε_i.

The asymptotic bias of the estimate under a small amount of contamination and certain regularity conditions is $B_i(z_{ki}^*) \approx \varepsilon_i \mathrm{IF}_i(z_{ki}^*, \psi)$. The sum of the expected values of $B_i^\mathsf{T}(z_{ki}^*)\, W B_i(z_{ki}^*)$, $i = 1, \ldots, N$, approximately equal to

$$\tilde{\Lambda}_{\chi,\varepsilon}(\psi) = \sum_{i=1}^{N} \varepsilon_i^2\, \mathbf{E}_{\chi_i}\left[\mathrm{IF}_i^\mathsf{T}(z_i, \psi)\, W\, \mathrm{IF}_i(z_i, \psi)\right], \tag{4}$$

where $\chi = (\chi_1, \ldots, \chi_N)^\mathsf{T}$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_N)^\mathsf{T}$, $\mathbf{E}_{\chi_i}$ is expection under the p.d.f. $\chi_i(z_i|\phi)$. This sum coincides with (2) for the weight function $s_i(z_i|\phi) = \varepsilon_i^2 \chi_i(z_i|\phi)$.

Let

$$\chi_i(z_i|\phi) = g_i^{1-\gamma}(z_i|\phi) / \mathrm{I}_{\gamma,i}(\phi), \tag{5}$$

where $\gamma \geq 0$ is a parameter, $\mathrm{I}_{\gamma,i}(\phi) = \int_{R^n} g_i^{1-\gamma}(z_i|\phi)\, d\mu$, and

$$\varepsilon_i = c\sqrt{\mathrm{I}_{\gamma,i}(\phi)\Delta_{\gamma,i}(\phi)},$$

where c is a proportionality factor, $\Delta_{\gamma,i}(\phi) = \left[\int_{R^n} g_i^{1+\gamma}(z_i|\phi)\, d\mu\right]^{\frac{\gamma}{1+\gamma}}$. Then weight function given by

$$s_i(z_i|\phi) = c^2 g_i^{1-\gamma}(z_i|\phi)\Delta_{\gamma,i}(\phi). \tag{6}$$

Optimum estimating function (3) is

$$\psi^*_{\gamma,i}(z_i,\phi) = C\left\{\frac{\partial}{\partial\phi}\ln g_i(z_i|\phi) + b_i\right\}\frac{g_i^{\gamma}(z_i|\phi)}{\Delta_{\gamma,i}(\phi)} \tag{7}$$

and corresponds to loss function

$$\rho_{\gamma,i}(\tilde{\zeta}_i,\phi) = -g_i^{\gamma}(\tilde{\zeta}_i|\phi)/\Delta_{\gamma,i}(\phi).$$

This loss function defines the MGDE with robustness parameter γ. Although, in our opinion, it is more natural to associate this estimator with the pseudo-spherical divergence (cf. [6, 9, 13]).

Distribution (5) has optimality properties from the point of view of information theory [3, 16]. In regression problems the amount of contamination can be considered as dependent on input variables (see also [6]).

If $g_i^{1-\gamma}(z_i|\phi)$ is not integrable, then we can consider some sequence of integrable functions $\eta_{i,u}(z_i|\phi)$, $u = 1, 2, \ldots$, converging to $g_i^{1-\gamma}(z_i|\phi)$, the correspondyng sequence of integrals $\Upsilon_{i,u}(\phi) = \int_{R^n}\eta_{i,u}(z_i|\phi)\, d\mu$, and sequence of amounts of contamination $\epsilon_{i,u} = c\sqrt{\Upsilon_{i,u}(\phi)\Delta_{\gamma,i}(\phi)}$. Then, under certain regularity conditions, the weight (6) will be the limit of functions $\epsilon_{i,u}^2\eta_{i,u}(z_i|\phi)/\Upsilon_{i,u}(\phi) = c^2\eta_{i,u}(z_i|\phi)\Delta_{\gamma,i}(\phi)$, and indicator (4) will be the limit of sequence of integrals $\tilde{\Lambda}_{\tilde{\eta}_u,\epsilon_u}(\psi)$, where $\tilde{\eta}_u = (\eta_{1,u}(z_i|\phi)/\Upsilon_{1,u}(\phi), \ldots, \eta_{N,u}(z_i|\phi)/\Upsilon_{N,u}(\phi))^{\mathsf{T}}$, $\epsilon_u = (\epsilon_{1,u}, \ldots, \epsilon_{N,u})^{\mathsf{T}}$ (see also [15]).

Note that we can postulate a constant amount of contamination and multiply the terms in sum (4) by the products $\mathrm{I}_{\gamma,i}(\phi)\Delta_{\gamma,i}(\phi)$ as additional weights of observations. Moreover, we obtain the estimating function (7), if we directly use weight (6) in the square of the weighted L_2-norm of the influence function (2) without using model of the point Bayesian contamination.

Choosing $s_i(z_i|\phi) = g_i^{1-\gamma}(z_i|\phi)$ and $s_i(z_i|\phi) = \chi_i(z_i|\phi)$ (χ_i from (5)) we obtain generalized radical estimators (GRE's) [3, 12, 14, 15, 16], but, in order to distinguish them, we will name the latter the Bayesian GRE (BGRE), since its weight function is p.d.f.

The GRE and BGRE don't have optimization formulation. Therefore, choosing the estimate as one of the estimating equation solutions can be problematic. In this case, a frequently used approach is to choose a suitable initial approximation with the calculation of the estimate using a local solution method. This initial approximation usually is a good (in some sense) estimate [3]. Such an estimate can be the MLE or the MGDE.

An alternative is the minimum beta-divergence (or density power divergence) estimator (MBDE) [11]. For non-homogeneous data, it is defined by the loss function [7]

$$\rho_i(\tilde{\zeta}_i, \phi) = \int\limits_{R^n} g_i^{1+\gamma}(z_i|\phi)\, d\mu - \left(1 + \gamma^{-1}\right) g_i^{\gamma}(\tilde{\zeta}_i|\phi).$$

The remarkable fact is that the MBDE applied to the qualitative response regression model is a version of the Bianco and Yohai estimator (cf. [2, 3]).

Note that the MGDE, GRE, and BGRE are actually the same for homogeneous data; for this case, optimality properties were discussed earlier (see, for example, [15, 16]).

2 Robust estimation of the ordered probit model

We will apply the above theory to regression with a discrete output variable (response), in this case μ is counting measure.

Let the assumed distribution of a discrete random variable ζ_i under the ith observation be given by a set of probabilities

$$\mathsf{P}\{\zeta_i = j|x_i, \phi\} = g(j|x_i, \phi),\ i = 1, \ldots, N,\ j = 1, \ldots, J,$$

where x_i is a vector of deterministic input variables.

For modeling dependence of the ordinal response on input variables the following cumulative link model is often used. Consider latent variable ζ_i^* under the ith observation that satisfies the regression model $\zeta_i^* = F(x_i)\alpha + e_i$, where $F(x_i)$ is a vector of nonconstant regressors (functions of input variables), α is a vector of parameters, e_i is a random error with a cumulative distribution function G. When $l_{j-1} < \zeta_i^* \le l_j$, where $-\infty \equiv l_0 < l_1 < \cdots < l_{J-1} < l_J \equiv +\infty$, the response ζ_i takes the value j. Corresponding probabilities are

$$g(j|x_i, \phi) = G\left(l_j - F(x_i)\alpha\right) - G\left(l_{j-1} - F(x_i)\alpha\right).$$

If errors have standard normal distribution, we obtain the ordered probit model. The vector of estimated parameters is $\phi = (l_1, \ldots, l_{J-1}, \alpha^{\mathsf{T}})^{\mathsf{T}}$.

Consider the following model: $J = 3$, single input variable (also it is regressor) takes values in $N = 301$ nodes of the uniform grid on the segment $X = [-1.5, 1.5]$, $\phi = (-1.5, 1.5, 2.8)^{\mathsf{T}}$.

The study of influence functions for the MGDE, MBDE, GRE, and BGRE shows that under small values γ all estimates are close. The MBDE differs from other estimators in that the influence of observations with values $j = 1$ and $j = 3$ increases in parts of X with small probabilities of these values as γ increases (all investigated in [3] estimators had similar properties), and in the case of $\gamma \ge 0.5$ such difference is too noticeable. In the case of $\gamma < 1$ other estimators are very close. In the case of $\gamma > 1$ the BGRE differs from other estimators in that the influence of observations with value $j = 2$ increases in parts of X with small probabilities of this value as γ increases. For this reason in the case of $\gamma \ge 1.3$ the BGRE seems too different from the other three estimators. This is due to the specificity of the BGRE weight function. The GRE and MGDE are close in the wide range of γ values. Thus, the MGDE seems to be the most preferable estimator.

Results are illustrated by Figure 1 where for the case of $\gamma = 1.5$ dependences of the element of the influence function corresponding to α on input variable are presented. Here and below, dependences for different values j are designated by lines of different styles: solid for $j = 1$, dotted for $j = 2$, dashed for $j = 3$. For comparing of influence functions it is convenient to use dependences $g(j|x, \phi)$; they are represented in Figure 2 by gray lines.

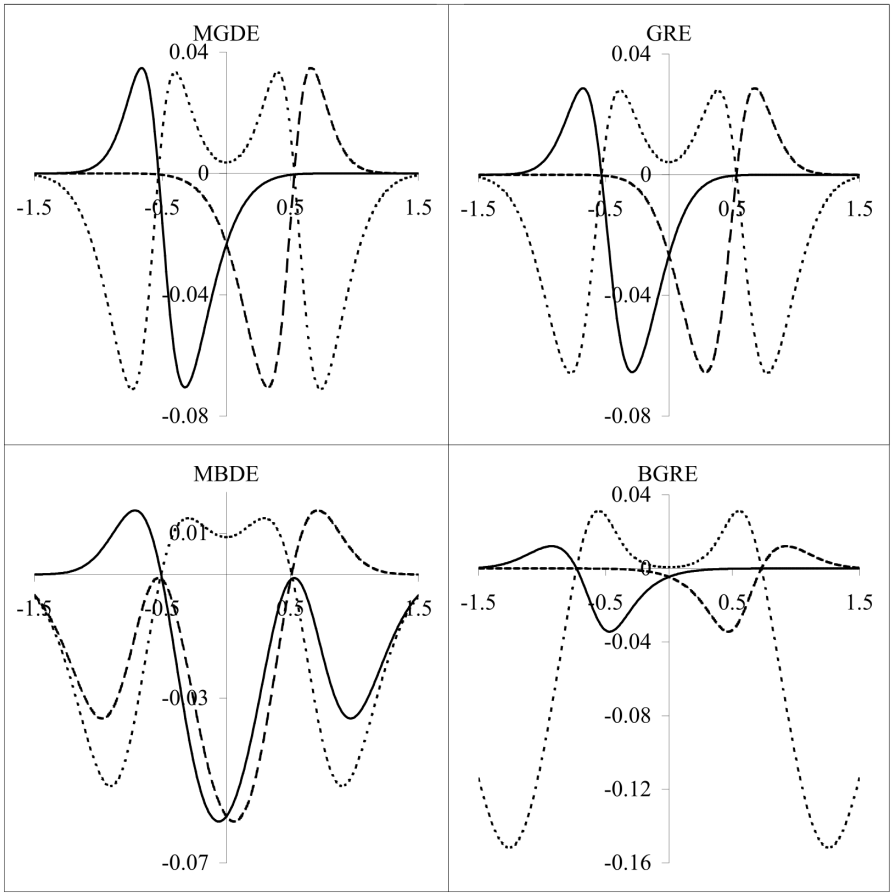


Figure 1: Influence functions

To evaluate the applicability of the MGDE, a small numerical experiment was performed using a simulated dataset. The MLE and MGDE were compared under the same model as above. Real distribution was homogeneous contaminated, contamination distribution was uniform, $\varepsilon = 0.1$. And the robustness parameter was $\gamma = 1$ (there is some analogy with the estimator of minimum variance sensitivity [17]). The results of the MLE are $\hat{\phi} = (-0.9808874, 0.7818217, 1.615199)^{\mathsf{T}}$. And corresponding results of the MGDE are $\hat{\phi} = (-1.444982, 1.253272, 2.317606)^{\mathsf{T}}$.

Figure 2 shows us estimated dependences of probabilities on the input variable; they are represented by black lines. The left panel corresponds to the MLE, the right panel corresponds to the MGDE. As a results of the experiment, the MGDE is less affected by contamination than the MLE.

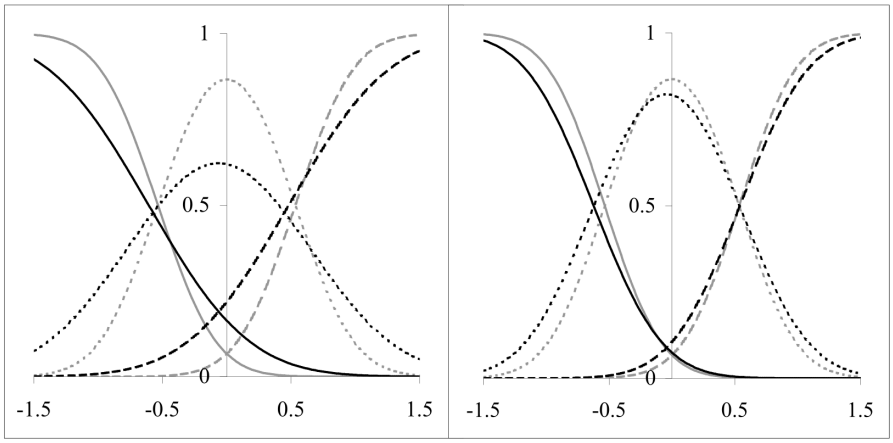


Figure 2: Ideal (gray lines) and estimated (black lines) probabilities

References

[1] Borovkov A.A. (1998). *Mathematical statistics*. Gordon and Breach, Amsterdam.

[2] Castilla E., Ghosh A., Martin N., Pardo L. (2018). New Robust Statistical Procedures for the Polytomous Logistic Regression Models. *Biometrics*. Vol. **74**, pp. 1282-1291.

[3] Chernikov Yu.V., Lisitsin D.V. (2017). Robust Polytomous Logistic Regression Based on Bianco and Yohai Estimator. *Proceedings of the International Workshop "Applied Methods of Statistical Analysis. Nonparametric Methods in Cybernetics and System Analysis"*. Novosibirsk, pp. 41-48.

[4] Croux C., Haesbroeck G., Ruwet C. (2013). Robust Estimation for Ordinal Regression. *Journal of Statistical Planning and Inference*. Vol. **143**, pp. 1486-1499.

[5] Dolgovykh E.M., Lisitsin D.V. (2015). Robust Estimation of Multivariate Regression Model in the Presence of Missing Data. *Proceedings of the International Workshop "Applied Methods of Statistical Analysis. Nonparametric Approach"*. Novosibirsk, pp. 64-71.

[6] Fujisawa H., Eguchi S. (2008). Robust Parameter Estimation with a Small Bias against Heavy Contamination. *Journal of Multivariate Analysis*. Vol. **99**, pp. 2053-2081.

[7] Ghosh A., Basu A. (2013). Robust Estimation for Independent Non-Homogeneous Observations using Density Power Divergence with Applications to Linear Regression. *Electronic Journal of Statistics*. Vol. **7**, pp. 2420-2456.

[8] Hampel F.R., Ronchetti E.M., Rousseeuw P.J., Stahel W.A. (1986). *Robust Statistics: The Approach Based on Influence Functions*. Wiley, New York.

[9] Hung H., Jou Z.-Y., Huang S.-Y. (2018). Robust Mislabel Logistic Regression without Modeling Mislabel Probabilities. *Biometrics*. Vol. **74**, pp. 145-154.

[10] Iannario M., Monti A.C., Piccolo D., Ronchetti E. (2017). Robust Inference for Ordinal Response Models. *Electronic Journal of Statistics*. Vol. **11**, pp. 3407-3445.

[11] Jones M.C., Hjort N.L., Harris I.R., Basu A. (2001). A Comparison of Related Density-Based Minimum Divergence Estimators. *Biometrika*. Vol. **88**, pp. 865-873.

[12] Kalinin A.A., Lisitsin D.V. (2011). Robust Estimation of Qualitative Response Regression Models. *Proceedings of the International Workshop "Applied Methods of Statistical Analysis. Simulations and Statistical Inference"*. Novosibirsk, pp. 303-309.

[13] Kanamori T. (2014). Scale-Invariant Divergences for Density Functions. *Entropy*. Vol. **16**, pp. 2611-2628.

[14] Lisitsin D.V., Formanchuk D.S. (2007). Robust Parameter Estimation for Heteroscedactic Regression. *Proceedings of the Russian Higher School Academy of Sciences*. No. **2**, pp. 69-75 (in Russian).

[15] Lisitsin D.V., Gavrilov K.V. (2011). About some properties of M-estimations. *Transaction of scientific papers of the Novosibirsk State Technical University*. No. **2**, pp. 61-68 (in Russian).

[16] Lisitsin D.V., Gavrilov K.V. (2016). Estimation of Distribution Parameters of a Bounded Random Variable Robust to Bound Disturbance. *Science Bulletin of the Novosibirsk State Technical University*. No. **2**, pp. 70-89 (in Russian).

[17] Shevlyakov G., Morgenthaler S., Shurygin A. (2008). Redescending M-estimators. *Journal of Statistical Planning and Inference*. V. **138**, pp. 2906-2917.

[18] Shurygin A.M. (1994). New Approach to Optimization of Stable Estimation. *Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach*. Kluwer, Dordrecht, pp. 315-340.

[19] Tabatabai M. A., et al. (2014). Robust Logistic and Probit Methods for Binary and Multinomial Regression. *Journal of Biometrics and Biostatistics*. Vol. **5**, issue 202.

Asymptotically efficient estimation of a drift coefficient in diffusion processes

Evgeny Pchelintsev and Svyatoslav Perelevskiy
Tomsk State University, Tomsk, Russia
e-mail: evgen-pch@yandex.ru, slavaperelevskiy@mail.ru

Abstract

The article studies the asymptotic properties of an adaptive model selection procedure for estimation an unknown drift coefficient in diffusion processes. It is shown that the procedure is asymptotically efficient, i.e. it is established that the asymptotic quadratic risk of the procedure coincides with the Pinsker constant, which provides an exact lower bound of the quadratic risk for all possible estimates.

Keywords: improved estimation, stochastic diffusion process, mean-square accuracy, oracle inequalities, Pinsker constant, asymptotic efficiency.

Introduction

Consider the problem of asymptotically efficient estimation of the unknown drift coefficient in diffusion process, described by the following stochastic differential equation:

$$\mathrm{d}y_t = S(y_t)\,\mathrm{d}t + \mathrm{d}w_t\,, \quad 0 \le t \le T\,, \tag{1}$$

where $(w_t)_{t\ge0}$ is a scalar standard Wiener process, the initial value y_0 is some given constant, and $S(\cdot)$ is an unknown function. Note that such models are widely used in financial markets, radio-physics, etc. [1]. The problem is to estimate the function $S(x)$, $x \in [a,b]$, from the observations $(y_t)_{0\le t\le T}$. The main goal of this paper to prove the asymptotic efficiency property of the improved model selection procedure proposed in [2] for estimating the function S in (1). The concept of asymptotic efficiency is associated with the optimal rate of convergence of the minimax risk, i.e. An important issue in the optimality results is the study of the exact asymptotic of the minimax risk. The problem of asymptotic non-parametric estimation in the model of heteroscedastic regression was studied by Efroimovich [3] and Pinsker [4]. To prove the asymptotic efficiency of the procedure, it is necessary to show that its asymptotic quadratic risk coincides with the lower bound defined by the Pinsker constant [5, 6]. In this paper, the problem is solved using an approach based on the model selection methods and sharp oracle inequalities. Recall that the model selection method appeared in the pioneering works of Akaike [7] and Mallows [8], in which proposed to introduce a penalization term in the criteria of maximum likelihood. Further, Barron, Birgé and Massart [9], Massart [10] and Kneip [11] developed this method to obtain non-asymptotic oracle inequalities in non-parametric regression models with Gaussian noise in discrete time. Unfortunately, this method cannot be applied in our case to prove an asymptotic efficiency property, since the coefficient

in main term of the resulting oracle inequalities is greater then one. For this reason, in this paper we will use the method proposed in [12]. This paper deals with the estimating the unknown function $S(x)$, $a \le x \le b$, in the sense of the mean square risk

$$\mathcal{R}(\widehat{S}_T, S) \,=\, \mathbf{E}_S \|\widehat{S}_T - S\|^2\,, \quad \|S\|^2 = \int_a^b S^2(x)\mathrm{d}x\,, \tag{2}$$

where $\widehat{S}_T$ is some estimate of S by observations $(y_t)_{0\le t\le T}$, $a < b$ are some real numbers. Here $\mathbf{E}_S$ is the expectation with respect to the distribution $\mathbf{P}_S$ of the random process $(y_t)_{0\le t\le T}$ given the drift function S. To obtain a reliable estimator of function S, it is necessary that the process (1) has the ergodicity property. For this we suppose that unknown function S belongs to the following functional class:

$$\Sigma_{L,N} = \{S \in Lip_L(\mathbb{R}) \,:\, |S(N)| \le L\,;\; \forall |x| \ge N,\; \exists\, \dot{S}(x) \in \mathbf{C}(\mathbb{R})$$
$$\text{such that} - L \le \inf_{|x|\ge N} \dot{S}(x) \le \sup_{|x|\ge N} \dot{S}(x) \le -1/L\}\,, \tag{3}$$

where $L > 1$, $N > |a| + |b|$, $\dot{S}(x)-$ derivative $S(x)$. For estimating the drift S in (1) Galtchouk and Pergamenshchikov [13] have proposed to apply the sequential approach. First step is a passage to a discrete time regression model by making use of the truncated sequential procedure introduced in [5]. To this end, at any point x_k of an equidistant partition of the interval $[a, b]$, we define a sequential procedure (τ_k, S_k^*) with a stopping rule τ_k and an estimators S_k^*. For $Y_k = S_k^*$ with $1 \le k \le n$, we come to the regression equation on some set $\Gamma \subseteq \Omega$ ($\sup_{S\in\Sigma_{L,N}} \mathbf{P}_S(\Gamma^c) \le \Pi_T$, where $\lim_{T\to\infty} T^m\,\Pi_T = 0$ for any $m > 0$):

$$Y_k = S(x_k) + \zeta_k\,. \tag{4}$$

Here, in contrast with the classical regression model, the noise sequence $(\zeta_k)_{1\le k\le n}$ has a complicated structure, namely,

$$\zeta_k = \sigma_k\,\xi_k + \delta_k\,, \tag{5}$$

where $(\sigma_k)_{1\le k\le n}$ is a sequence of some observed random variables, $(\delta_k)_{1\le k\le n}$ is a sequence of bounded random variables and $(\xi_k)_{1\le k\le n}$ is a sequence of i.i.d. random variables $\mathcal{N}(0,1)$ which are independent of $(\sigma_k)_{1\le k\le n}$.

In order to estimate the function S in model (4) we make use of the model selection method based on improved weighted least squares estimates proposed [18]. Improved estimation method in nonparametric regression models has been developed in [15, 16, 17].

1 Oracle inequalities

To estimate the unknown function in model (4), we use improved weighted least squares estimates, defined in [2],

$$S_\lambda^*(x_l) \,=\, \sum_{j=1}^{n} \lambda(j)\,\theta_{j,n}^*\,\phi_j(x_l)\,\mathbf{1}_\Gamma\,, \quad 1 \le l \le n\,, \tag{6}$$

where $(\phi_j)_{j\le 1}$ is an orthonormal functions system, the vector of weight coefficients $\lambda = (\lambda_1, ..., \lambda_n)$ belongs some finite set Λ from $[0,1]^n$,

$$\theta^*_{j,n} = \left(1 - \frac{c(d)}{\|\widetilde{\theta}_n\|}\mathbf{1}_{\{1\le j\le d\}}\right)\widehat{\theta}_{j,n}, \quad \|\widetilde{\theta}_n\|^2 = \sum_{j=1}^{d}\widehat{\theta}^2_{j,n}, \quad \widehat{\theta}_{j,n} = \frac{b-a}{n}\sum_{l=1}^{n} Y_l\phi_j(x_l)\,.$$

Here the coefficient $d \approx n^\epsilon/\ln n$, $0 < \epsilon < 1$, $c(d) \approx d/n$. Now we define the estimate for S in (1). We set for any $a \le x \le b$

$$S^*_\lambda(x) = S^*_\lambda(x_1)\mathbf{1}_{\{a\le x\le x_1\}} + \sum_{l=2}^{n} S^*_\lambda(x_l)\mathbf{1}_{\{x_{l-1}<x\le x_l\}}\,. \tag{7}$$

In order to obtain a good estimator, we have to write a rule to choose a weight vector $\lambda \in \Lambda$ in (7). It is obvious, that the best way is to minimize the empirical squared error with respect to λ:

$$\mathrm{Err}_n(\lambda) = \|S^*_\lambda - S\|^2_n \to \min\,.$$

Making use of (7) and the Fourier transformation of S imply

$$\mathrm{Err}_n(\lambda) = \sum_{j=1}^{n}\lambda^2(j)\theta^{*2}_{j,n} - 2\sum_{j=1}^{n}\lambda(j)\theta^*_{j,n}\,\theta_{j,n} + \sum_{j=1}^{n}\theta^2_{j,n}\,.$$

Since the coefficient $\theta_{j,n}$ is unknown, we need to replace the term $\theta^*_{j,n}\,\theta_{j,n}$ by some its estimator which we choose as

$$\widetilde{\vartheta}_{j,n} = \widehat{\theta}_{j,n}\theta^*_{j,n} - \frac{b-a}{n}s_{j,n} \quad \text{with} \quad s_{j,n} = \frac{b-a}{n}\sum_{l=1}^{n}\sigma^2_l\,\phi^2_j(x_l)\,.$$

One has to pay a penalty for this substitution in the empirical squared error. Finally, we define the cost function of the form

$$J_n(\lambda) = \sum_{j=1}^{n}\lambda^2(j)\theta^{*2}_{j,n} - 2\sum_{j=1}^{n}\lambda(j)\,\widetilde{\vartheta}_{j,n} + \rho P_n(\lambda)\,,$$

where the penalty term is defined as

$$P_n(\lambda) = \frac{b-a}{n}\sum_{j=1}^{n}\lambda^2(j)s_{j,n}$$

and $0 < \rho < 1$ is some positive constant which will be chosen later. We set

$$\widehat{\lambda} = \mathrm{argmin}_{\lambda\in\Lambda}\,J_n(\lambda)$$

and define an estimator of S of the form (7):

$$S^*(x) = S^*_{\widehat{\lambda}}(x) \quad \text{for} \quad a \le x \le b\,. \tag{8}$$

Now we obtain the non asymptotic upper bound for the quadratical risk of the estimator (8).

Theorem 1. *Let $\Lambda \subset [0,1]^n$ be any finite set such that the first $d \le n$ components of the weight vector λ are equal to 1. Then, for any $n \ge 3$ and $0 < \rho < 1/6$, the estimator (8) satisfies the following oracle inequality*

$$\mathbf{E}_S\|S^* - S\|_n^2 \le \frac{1+6\rho}{1-6\rho} \min_{\lambda\in\Lambda} \mathbf{E}_S\|\widehat{S}_\lambda - S\|_n^2 + \frac{\Psi_n(\rho)}{n}\,,$$

where $\lim_{n\to\infty} \Psi_n(\rho)/n = 0$.

Now we consider the estimation problem (1) via model (4). We apply the estimating procedure (8) with special weight set introduced in [5] to the regression scheme (4). Denoting $S^*_\alpha = S^*_{\lambda_\alpha}$ we set

$$S^* = S^*_{\widehat{\alpha}} \quad \text{with} \quad \widehat{\alpha} = \operatorname{argmin}_{\alpha\in\mathcal{A}_\varepsilon} J_n(\lambda_\alpha)\,.$$

Theorem 2. *Assume that $S \in \Sigma_{L,N}$ and the number of the points $n = n(T)$ in the model (4) . Then the procedure S^* satisfies, for any $T \ge 32$, the following inequality*

$$\mathcal{R}(S^*, S) \le \frac{(1+\rho)^2(1+6\rho)}{1-6\rho} \min_{\alpha\in\mathcal{A}_\varepsilon} \mathcal{R}(S^*_\alpha, S) + \frac{\mathcal{B}_T(\rho)}{n}\,,$$

where $\lim_{T\to\infty} \mathcal{B}_T(\rho)/n(T) = 0$.

2 Asymptotic efficiency

In order to study the asymptotic efficiency we define the following functional Sobolev ball

$$W^0_{k,r} = \{f \in \mathbf{C}^k_0([a,b]) \,:\, \sum_{j=0}^{k} \|f^{(j)}\|^2 \le r\}\,, \tag{9}$$

where $r > 0$ and $k \ge 1$ are some unknown parameters, $\mathbf{C}^k_0([a,b])$ is the space of k times differentiable functions $f : \mathbb{R} \to \mathbb{R}$ such that

$$f^{(i)}(x) = 0\,, \quad \text{for} \quad 0 \le i \le k-1 \quad \text{and} \quad x \notin [a,b]\,.$$

We will call such functions *periodic on the interval* $[a,b]$. Let S_0 be a fixed $k+1$ times continuously differentiable function from $\Sigma_{L,N}$. We set

$$\Theta_{k,r} = \{S = S_0 + f\,,\, f \in W^0_{k,r}\}\,. \tag{10}$$

In order to formulate our asymptotic results we define the following normalizing coefficient

$$\gamma(S) \,=\, ((1+2k)r)^{1/(2k+1)} \left(\frac{J(S)k}{\pi(k+1)}\right)^{2k/(2k+1)} \tag{11}$$

with

$$J(S) = \int_a^b \frac{1}{q_S(x)}\mathrm{d}x\,, \quad q_S(x) \,=\, \frac{\exp\{2\int_0^x S(z)\mathrm{d}z\}}{\int_{-\infty}^{+\infty}\exp\{2\int_0^y S(z)\mathrm{d}z\}\mathrm{d}y}.$$

It is well known that for any $S \in \Theta_{k,r}$ the optimal rate of convergence is $T^{-2k/(2k+1)}$ (see, for example, [18]). On the basis of the model selection procedure (8) in the next section we construct the adaptive procedure S^* for which we obtain the following asymptotic upper bound for the quadratic risk.

Theorem 3. *The quadratic risk* (2) *for the estimating procedure* S^* *has the following asymptotic upper bound*

$$\limsup_{T\to\infty} T^{2k/(2k+1)} \sup_{S\in\Theta_{k,r}} \frac{\mathcal{R}(S^*,S)}{\gamma(S)} \le 1\,. \tag{12}$$

Moreover, we show that this upper bound is sharp in the following sense.

Theorem 4. *For any estimator* $\widehat{S}$ *of* S *measurable with respect to* $\mathcal{F}_T^y$,

$$\liminf_{T\to\infty} \inf_{\widehat{S}} T^{2k/(2k+1)} \sup_{S\in\Theta_{k,r}} \frac{\mathcal{R}(\widehat{S},S)}{\gamma(S)} \ge 1\,, \tag{13}$$

where $\mathcal{F}_T^y$ *is a* $\sigma-$*field generated by observations* $(y_t)_{0\le t\le T}$.

Remark 1. *It should be noted that the choice of the functional class* $\Theta_{k,r}$ *in the form of* (10) *is related to the ergodicity of the process* (1). *This property is provided when the drift derivative is negative on the outside of a finite interval. The last excludes the choice of periodic functions as a class of admissible drifts. For this reason, we use the Sobolev ball of periodic functions with a non periodic center* S_0 *as a class of admissible drift functions.*

Remark 2. *Note that the inequalities* (14) *and* (13) *imply that the function* (11) *is the Pinsker constant in this case (cf. [4]).*

Corollary 1. *From Theorems 2 and 3 it follows that the procedure for choosing a model* S^*, *defined in* (8), *is asymptotically efficient, i.e.*

$$\lim_{T\to\infty} T^{2k/(2k+1)} \sup_{S\in\Theta_{k,r}} \frac{\mathcal{R}(S^*,S)}{\gamma(S)} = 1\,. \tag{14}$$

3 Numerical simulations

We suppose that in the model (1)

$$S(x) = x^2\sin(2\pi x) + x^2(1-x)\cos(4\pi x).$$

For weight coefficients we choose $n = T$,

$$k^* = 100 + \sqrt{\ln n}, \quad \varepsilon = \frac{1}{\ln n}, \quad m = \ln^2 n, \quad \omega_\alpha = 100 + (A_\beta t n)^{\frac{1}{2\beta+1}}\,.$$

The empirical risk:

$$\mathcal{R}(S^*,S) = \frac{1}{1000}\sum_{m=1}^{1000} \|S_m^* - S\|_n^2.$$

Table 1 shows the results of the behavior of empirical mean-square risks for the proposed estimation procedure (8).

Table 1: Empirical quadratic asymptotic risks

n	501	1001	2001	10001
$T^{2k/(2k+1)}\frac{\mathcal{R}(S^*,S)}{\gamma(S)})$	4.7257	2.0856	1.0072	0.9012

From Table 1 it is clear that with an increase in the number of observations n, the normalized empirical mean-square risks tend to unity, which confirm numerically the Corollary 1.

The figures show the behavior of observation processes $(y_t)_{0\leq t\leq 1}$, function S (red line), and improved estimate S^* (green line):

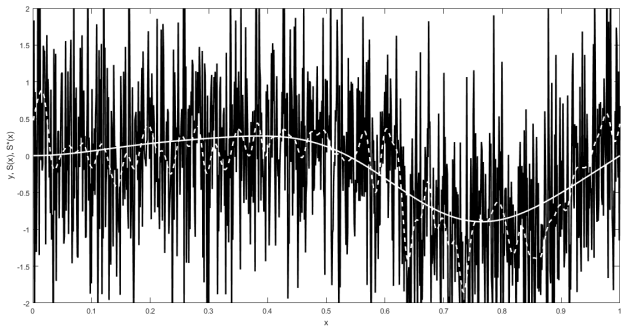


Figure 1: n=501

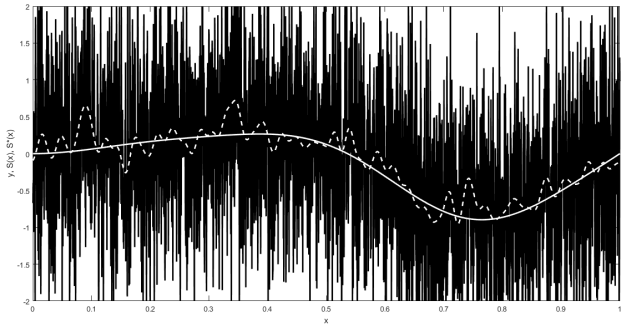


Figure 2: n=1001

Acknowledgements

This work was supported by the RSF grant no 17-11-01049.

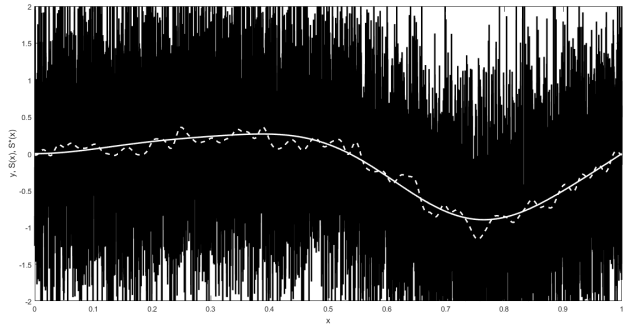


Figure 3: n=10001

References

[1] Kutoyants Yu. (2004). *Statistical inference for ergodic diffusion processes.* Springer-Verlag, London.

[2] Pchelintsev E.A., Perelevskiy S.S., Makarova I.A. (2018). Improved nonparametric estimation of the drift in diffusion processes. *Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-Matematicheskie Nauki.* Vol. **160**, No **2**, pp. 364–372.

[3] Efroimovich S. (2007). Sequential design and estimation in heteroscedastic nonparametric regression. *Sequential Analysis.* No **26**, pp. 3–25. Springer, New York, 1998.

[4] Pinsker M.S. (1981) Optimal filtration of square integrable signals in gaussian white noise. *Problems Transimis. information.* No **17**, pp. 120–133.

[5] Galtchouk L., Pergamenshchikov S. (2006) Asymptotically efficient sequential kernel estimates of the drift coefficient in ergodic diffusion processes. *Statistical Inference for Stochastic Process.* Vol. **9**, No. **1**, pp. 1–16.

[6] Ibragimov I.A., Khasminskii R.Z. (1981). *Statistical Estimation: Asymptotic Theory.* Springer, New York.

[7] Akaike H. (1974). A new look at the statistical model identification. *IEEE Trans. on Automatic Control.* Vol. **19**, No. **7**, pp. 716–723.

[8] Mallows C. (1973). Some comments on C_p. *Technometrics.* Vol. **15**, pp. 661–675.

[9] Barron A., Birgé L., Massart P. (1999). Risk bounds for model selection via penalization. *Probab. Theory Relat. Fields.* Vol. **113**, pp. 301–415.

[10] Massart P. (2005). *A non-asymptotic theory for model selection.* European Congress of Mathematics. Eur. Math. Soc., Zurich.

[11] Kneip A. (1994). Ordered linear smoothers. *Annals of Statistics*. Vol. **22**, pp. 835–866.

[12] Galtchouk L.I., Pergamenshchikov S.M. (2009). Adaptive asymptotically efficient estimation in heteroscedastic nonparametric regression. *J. Korean Stat. Soc.* Vol. **38**, No **4**, pp. 305–322.

[13] Galtchouk L.I., Pergamenshchikov S.M. (2011). Adaptive sequential estimation for ergodic diffusion processes in quadratic metric. *Journal of Nonparametric Statistics*. Vol. **23**, No **2**, pp. 255–285.

[14] Pchelintsev E., Perelevskiy S. (2019). Adaptive estimation in heteroscedastic nonparametric regression. *Tomsk State University Journal of Mathematics and Mechanics*. No. **57**, pp. 38–52.

[15] Pchelintsev E.A., Pergamenshchikov S.M. (2018). Oracle inequalities for the stochastic differential equations. *Statistical Inference for Stochastic Processes*. Vol. **21**, No **2**, pp. 469–483.

[16] Pchelintsev E., Pchelintsev V., Pergamenshchikov S. (2018). Non asymptotic sharp oracle inequalities for the improved model selection procedures for the adaptive nonparametric signal estimation problem. *Communications - Scientific Letters of the University of Zilina*. Vol. **20**, No **1**, pp. 73–77.

[17] Pchelintsev E. A., Pergamenshchikov S. M. (2019). Improved model selection Method for an adaptive estimation in semimartingale regression models. *Tomsk State University Journal of Mathematics and Mechanics*. Vol. **58**, pp. 14–31.

[18] Galtchouk L.I., Pergamenshchikov S.M. (2004). Nonparametric sequential estimation of the drift in diffusion via model selection. *Mathematical Methods of Statistics*. Vol. **13**, pp. 25–49.

On controlled processes of multidimensional discrete-continuous systems

A. V. Medvedev
Siberian Federal University, Krasnoyrsk, Russian Federation
e-mail: mav2745@mail.ru

Abstract

When solving real-world problems, there are often processes that take place not in the area defined by the vectors of input and output variables, but in a sub-area. The article describes the following types of processes: H-processes, K-processes.

Keywords: Discrete-continuous systems, H-processes, K-processes.

Introdution

The controlled processes that occur continuously in time and the control variables that are carried out at discrete points in time traditionally refers to the class of discrete-continuous processes. In particular, A.A. Feldbaum in the first works on the theory of dual control in the monograph [2] considered the following system.

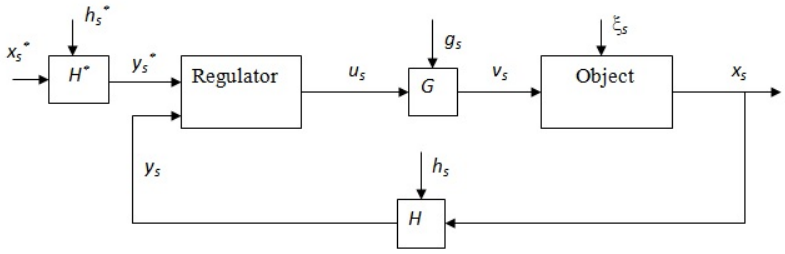


Figure 1: Block diagram of the discrete-continuous control system

In Figure 1, the following symbols are used: x_s^* is setting effect, that through the channel H^* is mixed with the noise h_s^* and enters as a regulator y_s^*; x_s is an output of the object, that passing through channel H and mixing with the noise h_s in the form of y_s also enters the regulator; u_s is a control action that, passing through the channel G and mixing with the g_s interference, comes in the form of ν_s to the object that is affected by the interference ξ_s; s is discrete time. The object was represented as $x_s = \alpha\nu_s$, i.e., the object was parameterized with the accuracy of the parameter α, and the equations of communication channels were assumed as additive ones with normally distributed noise.

When identifying dynamic processes, a situation often arises when the measurement of output variables is carried out in considerable time intervals. Moreover, the technology of these variables measuring is such that the time spent on it can be quite significant. As a result, the total time exceeds the object constant. Ultimately, it

leads to the fact that we are forced to consider the object as inertial with delay. Such a process on the appropriate channel of the object should be presented in the form

$$x(t) = f(u(t-\tau), \xi(t)), \tag{1}$$

where x(t) is the output variable of the object; $u(t-\tau)$ is input variable, here τ is delay; ξ is random perturbation influencing the object; (t) is continuous time period.

The measurement channels with interferences h^u, h^x, and measurement discreteness $\Delta T >> \Delta t$ are given in Figure 1. In principle, the object delay τ may be absent and then we deal with a normal dynamic object, but due to the large value of ΔT. In this case we are to consider it as a static delay determined by the measurement time duration of the variable x(t). Note that the time constant of the object is much less than ΔT, i.e., $\Delta T >> \Delta t$. Thus, a sample of observations in a discrete form can be represented as follows: $u[t], x[t+n+m]$, where $t = 1, 2, \ldots, s$; n is delay discreteness, $n = \tau/\Delta t$, m is delay caused by the duration of the monitoring, $m = \Delta T/\Delta t$. Carrying out the implementation shift in cycles, a sample of observations can be rewritten as follows: $\{u_t, x_t, t = \overline{1,s}\}$ without reducing the generality, reduce the task to identification of a static object.

Consider $u = (u1, \ldots, u_k) \in \Omega(u) \subset R^k, x \in \Omega(x) \subset R^1$. Generally speaking, each component of the vector $u_i \in [a_i; b_i], i = \overline{1,k}$, and $x \in [c; d]$. In real processes investigation, values of the coefficients $\{a_i, b_i, cd\}, i = \overline{1,k}$, are always known. In technological processes, values of these coefficients are controlled by the technological regulations (by a technological map). Further, we will take these intervals as single ones without breaking the generality, then $\Omega(u)$ is a single hypercube, and $\Omega(x) = [0; 1]$, i. e, $x \in [0; 1]$.

Usually the identification problem of the static object is reduced to parametric identification [3, 3, 4], that consists of two main stages: selection (definition) of the parametric model (1) in the form $\hat{x} = \hat{f}(u, \alpha)$, where α is a parameter vector, and the subsequent estimation of the parameters α based on the incoming sample elements $(u_1, x_1), (u_2, x_2), \ldots, (u_s, x_s)$, i.e., getting the estimation α_s. In this case, an adaptive model will be as follows:

$$\hat{x}_s = \hat{f}(u, \alpha_s) \tag{2}$$

If a row is taken as a function $\hat{f}(u, \alpha)$

$$\hat{x} = \hat{f}(u, \alpha) = \sum_{j=1}^{N} \alpha_j \varphi_j(u), \tag{3}$$

where $\varphi_j(u), j = \overline{1,N}$ is a system of linearly independent functions of a vector argument $u = (u_1, \ldots, u_k) \in \Omega(u)$, then, according to the method of stochastic approximations [3], we obtain

$$\alpha_s^l = \alpha_{s-1}^l + \gamma_s^l (x_s - \sum_{j=1}^{N} \alpha_{s-1}^l \varphi_l(u_s)) \varphi_l(u_s), l = 1, \ldots, N, \tag{4}$$

where $\gamma_s^l, l = \overline{1,N}$ are Robbins – Monroe coefficients.

This is the general scheme for solving parametric identification problems. Note that the weakness here is a choice of the parametric structure of the model. If a rather serious mistake has been made at the first stage, then the resulting model is unlikely to be satisfactory. Note that the models of the form (2) are hypersurfaces in the object input-output variable region: $(u, x) \in \Omega(u, x) \subset R^{k+1}$.

Let us analyze two important circumstances that arise while modeling real processes. The first of these is that a sample size s $\{x_i, u_i, i = \overline{1, s}\}$ is often rather insufficient relative to the dimension of the vector $u = (u_1, \ldots, u_k) \in \Omega(u)$, as mathematical statistics requires. For example, in practice we often have the following situation when $20 \ldots 30$, and $s = 900 \ldots 1\,000$, and therefore a satisfactory solution to the problem of identification cannot be obtained. The second circumstance is that if, according to the available data, a model of the type (2) has been developed, and then at $u \in \Omega(u) \subset R^k$ we can get the estimation $x_s \notin \Omega(x)$, i.e., the estimation of x outside the process regulations and even physically unrealizable values x(u). Both of these circumstances can be explained taking into account the following considerations.

So, the investigated process without loss of generality proceeds in a single cube $\Omega(u, x) = \Omega(u_1, u_2, x) \subset R^3$. If we omit the influence of random perturbations $\xi(t)$ and measurement mistakes u_1, u_2, x, i.e., the absence of h^u, h^x and ξ (Figure 1) then, for simplicity, the process takes place in the space $\Omega^H(u, x) \subset \Omega(u, x)$, as follows from the model of the form (2), which is a surface in the space $\Omega(u, x)$ (Figure 2).

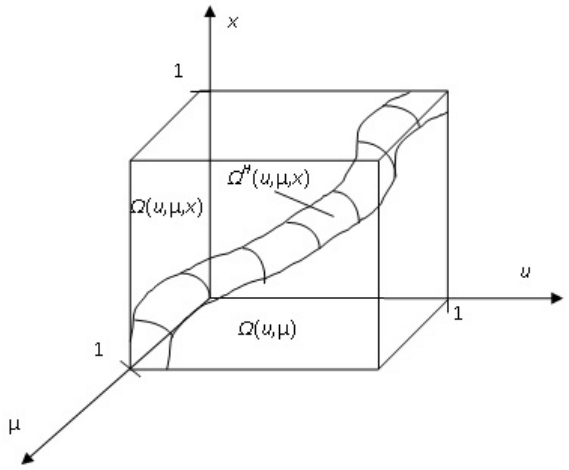


Figure 2: Scheme of the H-process

The actual values that process variables will take may be known to the researcher, for example, from process regulations. Thus, a real process takes place in a certain region $\Omega(u_1, u_2, x)$, in our example, in a unit cube. In the case of the stochastic dependence of the input variables, a process does not proceed in the entire region $\Omega(u_1, u_2, x)$, but only in its subspace $\Omega^H(u_1, u_2, x) \subset \Omega(u_1, u_2, x)$ which is always unknown. And since we do not know anything about the subspace $\Omega^H(u, x)$, we cannot say for sure that the investigated object has this feature. This is the main problem of modeling this kind of processes, which is called H-processes. When one

uses models of the form (2) for individual values of input variables of the process $u \in \Omega(u) \subset R^m$, for which a relationship $u \notin \Omega^H$ holds, we can get estimation $\hat{x_s} \notin \Omega^H(u,x)$ (point B in Figure 2) or even $\hat{x_s} \notin \Omega(u,x)$ (point C in Figure 2). In this case, the identification results will be unsatisfactory.

If the investigated process has a tube structure, then models (2) and (3) need to be corrected as follow:

$$\hat{x_s}(u) = I_s(u)\hat{f}(u,\alpha_s), \tag{5}$$

or

$$\hat{x_s}(u) = I_s(u)\sum_{j=1}^{N}\alpha_j\varphi_j(u), \tag{6}$$

where the indicator $I_s(u)$ is as follows:

$$I_s(u) = \begin{cases} 1, & \text{if } u \in \Omega_s^H(u), \\ 0, & \text{if } u \notin \Omega_s^H(u) \end{cases} \tag{7}$$

Models (5) and (6) will be called H-models.

Generally speaking, it should be noted that the space $\Omega^H(u)$, is unknown and only a sample $\{x_i, u_i, i = \overline{1,s}\}$ is known. If an indicator is equal to zero, then the estimation $\hat{x}(u)$ cannot be calculated, i.e., with such values of component vector $u \in \Omega(u)$ the process cannot proceed. If an indicator $I_s(u)$ at any value $u \in \Omega(u)$ is equal to one, then a model (5) coincides with model (2), and model (6) coincides with model (3).

As the estimation of the indicator $I_s(u)$, it is possible to take the following approximation [5]:

$$I_s(u) = sgn(sc_s)^{-1}\sum_{i=1}^{s}\Phi(c_s^{-1}(x_s(u) - x_i))\prod_{j=1}^{k}\Phi(c_s^{-1}(u^j - u_i^j)), \tag{8}$$

where

$$x_s(u) = \sum_{i=1}^{s}x_i\prod_{j=1}^{k}\Phi(c_s^{-1}(u^j - u_i^j))/\sum_{i=1}^{s}x_i\prod_{j=1}^{k}\Phi(c_s^{-1}(u^j - u_i^j)) \tag{9}$$

a blur parameter c_s and the bell-shaped function $\Phi(\cdot)$ satisfy the convergence conditions given in [5].

Thus, at a known value $u = u' \in \Omega(u)$, first it is necessary to build the estimation $x_s(u = u')$ using the formula (9), and then it is necessary to calculate the indicator and only then to use the models (5) or (6) if the indicator turns out to be equal to zero. If the indicator is equal to one, then this means that although $u \in \Omega(u)$ but $u \notin \Omega^H(u)$, i.e., vector components $u = u' = (u'_1, \ldots, u'_k)$ are defined incorrectly and the actual flowing tube process does not correspond to the set of specified values of the vector component $u = u'$. It is natural to assume that the identification process of the object in a parametric formulation should also be carried out taking into account

the tube structure of the object. In conclusion, we note that the nature of H-processes differs from fractal [6] and attractors.

If we interpret H-processes in a more general case as a function of several variables, then the variability of this function over time can be shown in the following chain of ratios operating in time:

The diagram of a dynamic object and the control of input and output variables is given below.

The input of the investigated process receives a vector of input variables $u(t) = (u_1(t), \dots u_n(t)) \in \Omega(u) \subset R^n$, the a vector of input variables gives $x(t) = (x_1(t), \dots, x_m(t)) \in \Omega(x) \subset R^m, \hat{x}(t)$ is the output of the model. Both variables are controlled at discrete instants of time through the interval Δt. In the process of the object investigation, a sample of observations $x_i = (x_{i1}, \dots, x_{im}), u_i = (u_{i1}, \dots, u_{in}), i = \overline{1, s}$ (where s is a training sample) can be obtained $\xi(t)$ is a vector of random impacts operating on an object whose expectation is equal to zero, and the dispersion is limited. The random noises $h^u(t)$ and operating in the measurement channels also have zero expectation and limited dispersion.

The peculiarity of identifying a multidimensional object is that the investigated process is described by a system of implicit stochastic equations.

$$F_j(u(t-\tau), x(t), \xi(t)) = 0, j = \overline{1, m} \tag{10}$$

where $F_j(\cdot)$ are unknown, τ is a delay along the various channels of the multidimensional system. Here τ means a delay known for all channels of the process under study. Further, τ will be omitted for reasons of simplicity. Objects will be treated as static. In subsequent models one can easily achieve it by the corresponding shift of the elements of the sample of observations $x_i = (x_{i1}, \dots, x_{im}), u_i = (u_{i1}, \dots, u_{in}), i = \overline{1, s}$. The identification problem is to build a model of the system presented in Figure 3 with the presence of a priori information and the training sample $x_i = (x_{i1}, \dots, x_{im}), u_i = (u_{i1}, \dots, u_{in}), i = \overline{1, s}$. When identifying multidimensional inertialess systems, the representation of the object model is traditionally used in the form:

$$\hat{x_j} = f_j(u_1, u_2, \dots, u_n, \alpha), j = \overline{1, m} \tag{11}$$

where the parametric structure of the object is found from the available a priori information, and α is a vector of parameters estimated by the existing or incoming training sample.

In the case of a stochastic components dependence of the output variables of an object, its model in the parametric version is:

$$\hat{F}_j(u, x, \alpha) = 0, j = \overline{1, m} \tag{12}$$

From practical considerations, it is natural to assume that at $x(t) = (x_1(t), \dots, x_m(t)) \in \Omega(x) \subset R^m$ and $u(t) = (u_1(t), \dots, u_n(t)) \in \Omega(u) \subset R^n$ a system has a unique root.

Due to the lack of a priori information, the type of the functions $\hat{F}_j(\cdot), j = \overline{1, m}$ cannot be determined with accuracy to the parameters α. It leads to the necessity

to consider the investigated process as a T-process, and its model as a T-model respectively.

In general, the investigated multidimensional system that implements the T-process can be represented in Figure 3.

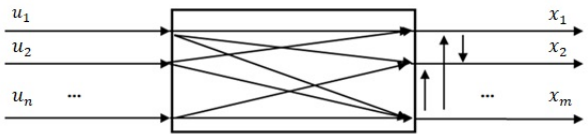


Figure 3: Multidimensional object

In the figure, the following symbols are taken: $u = (u_1, \ldots, u_n)$ is the n-dimensional vector of input variables, $x = (x_1, \ldots, x_m)$ is the m-dimensional vector of the output variables. Through various channels of the investigated process, the dependence of the j-th component of the vector u can be represented as a dependence on certain components of the vector u: $x^{<j>} = f_j(u^{<j>}), j = (\overline{1,m})$. Each j-th channel depends on several components of the vector u, for example $u^{<5>} = (u_1, u_3, u_6)$ where $u^{<5>}$ means a compound vector. The vector components $\vec{u} : u_2, u_4$ are not included in its structure due to the reasons of the nature of the investigated process, or $x^{<4>} = (x_1, x_3)$ is also a compound vector. The vertical arrows in Figure 3 for the components of the vector x(t) illustrate their stochastic dependence, which is unknown. In this case, the T-model of such a process should be considered as a system:

$$\hat{F}_j(u^{<j>}, x^{<j>}) = 0, j = \overline{1,m} \tag{13}$$

As a result of measuring input and output variables, a sample $x_i = (x_{i1}, \ldots, x_{im})$, $u_i = (u_{i1} \ldots, u_{in}), i = (\overline{1,s})$, used to develop an adaptive model of this object, can be obtained. In this case it is necessary to solve a system of the type (13) for the given values of input impacts $u_1 = u'_1, u_2 = u'_2, \ldots, u_n = u'_n$. As a result, we can obtain estimates of the components of the vector x for the corresponding values of the input impact.

K-models. The diagram that takes into account the reality while investigating some technological processes is given below.

Consider the problem of building a model of a dynamic process, shown in Figure 4. Note that the time intervals ΔT and T significantly exceed a constant of the object time along all other channels. Without breaking the generality, one can assume that the control of variables $u(t), \mu(t), \omega(t), x(t)$ is carried out in time interval $\Delta t, \Delta t << \Delta T << T$. Consequently, the processes for channels q(t) and z(t) can be assigned to the class of inertialess with a delay, and through the channels $\omega(t)$ and $x(t)$ can be assigned to the dynamic class, since the control of variables $u(t), \mu(t), \omega(t)$ is carried out in an interval Δt that is significantly smaller than the constant of the object time through the corresponding channels. In this case, a rather general K-model can be accepted as

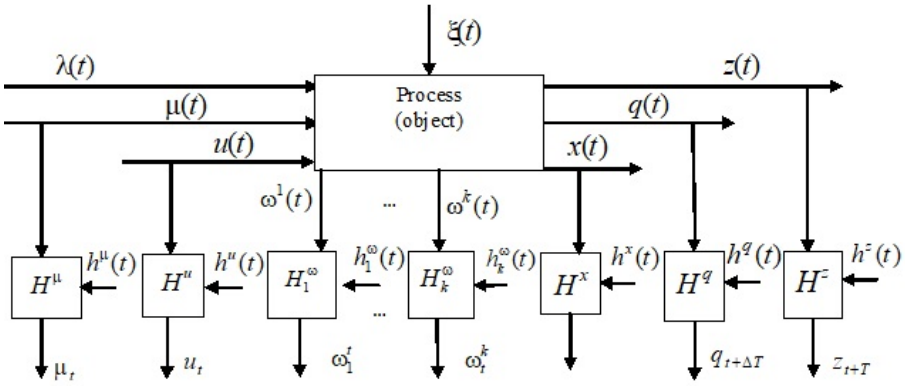


Figure 4: Block diagram of the interrelated stochastic process

$$\begin{cases} \hat{f}_i(u^{<i>}(t-\tau), \mu^{<i>}(t-\tau), \omega^{<i>}(t-\tau), x^{<i>}(t), \frac{dx^{<i>}(t)}{dt}, \\ \frac{d^2x^{<i>}(t)}{dt^2}, \ldots, \alpha) = 0, i = \overline{1,k}, \\ \hat{f}_i(u^{<i>}(t-\tau), \mu^{<i>}(t-\tau), \omega^{<i>}(t-\tau), x^{<i>}(t), q^{<i>}(t)z^{<i>}(t), \beta)I_s^i = 0, \\ i = \overline{k+1,l}, \\ \hat{S}_i(u^{<i>}(t-\tau), \mu^{<i>}(t-\tau), \omega^{<i>}(t-\tau), x^{<i>}(t), q^{<i>}(t)z^{<i>}(t), W_s^{<i>}) = 0, \\ i = \overline{l+1,m}, \end{cases} \tag{14}$$

Let us explain the symbols of the vector $x^{<i>}(t)$, which is a compound vector, in particular $x^{<1>}(t) = (x_1(t), x_3(t)), x^{<2>}(t) = (x_2(t), x_3(t), x_4(t))$, and etc. This vector is not composed of all the components of the vector , but of their part, due to the nature of the investigated process. Other variables $u^{<i>}(t-\tau), \mu^{<i>}(t-\tau), \omega^{<i>}(t-\tau)$ are also the corresponding compound vectors. Here it is important to pay attention to the fact that various compound vectors are formed exclusively from practical considerations when analyzing various real processes.

The first group of equations system (14) is found on the basis of the known fundamental laws that correspond to the process under study with an accuracy of parameters α. The second group of equations of the object is obtained on the basis of the available prior information up to the vector of parameters β. The third group of equations $\hat{S}_i$ is not known up to parameters, but a class of functions describing the relationship between input-output and intermediate variables is determined on the basis of a priori information. A symbol $W_s^{<i>}$ appearing here is a combination of all the i-th observations of variables with volume s:

$$W_s^{<i>} = (u^{<i>}, \mu^{<i>}, \omega^{<i>}, x^{<i>}, q^{<i>}z^{<i>}), i = \overline{l+1,m} \tag{15}$$

The values estimation of the components of the vectors of output variables

$x(t), q(t), z(t)$ can be found as a result of solving the system of equations (14) with the fixed values $u(t), \mu(t), \omega(t)$.

K-models are fundamentally different from those generally accepted first of all because they take into account all the available variables and interrelations between them in a situation where the discreteness of control of the latter is significantly different, as are the levels of a priori information about the various channels of the process being investigated. Thus, K-models are an organic synthesis that describes the process under investigation or a system of interconnected objects in all their diversity.

Thus, various processes occurring in real-life processes are considered. The following notion should be noted. H-T-processes appear as a result of the real processes functioning. It generates H-T-models. And K-models are obtained as a result of a lack of a priori information through various channels and they are based on the triad: fundamental laws, parametric equations obtained as a result of engineering research and some dependencies due to the lack of a priori information when only qualitative properties are known. It should also be noted that the exact mathematical formulation of the control problem in Figure 1, taken from the monograph written by A.A. Feldbaum, is a removed a bit from the specific operating systems. The representation of the object in Figure 4 does not make possible to get a similar mathematical formulation of the problem due to the lack of a priori information through various channels, and it means that a more adequate formulation of the control problem reflecting the reality is needed.

References

[1] Feldbaum A. A. Fundamentals of the theory of optimal automatic systems. - M.: Fizmatgiz, 1963. - 553 p.

[2] Tsypkin Ya. Z. Adaptation and training in automatic systems. - M.: Science, 1968. - 399 p.

[3] Aikhoff P. Fundamentals of identification of control systems. - M.: Mir, 1975. - 680 p.

[4] Ljung L. Systems identification. - M.: Science, 1991. - 432 p.

[5] Medvedev A.V. Fundamentals of the theory of nonparametric systems. - Krasnoyarsk: from-on SibGAU, 2018. - 727 p.

[6] Mandelbrot B. The Fractal Geometry of Nature / Izhev. Institute of computer researches; Nauch.-issled. center "Regular and chaotic dynamics." - Moscow: Izhevsk, 2010. - 655 p.

On levels of a priori information in the of identification and control problems

A. V. Medvedev
Siberian Federal University, Krasnoyrsk, Russian Federation
e-mail: mav2745@mail.ru

Abstract

In the formulation of the problem of identification and control, the importance of the available volume of a priori information about the process under study. A priori information is classified by the levels of a priori information depending on the volume. The non-parametric level of a priori information has been reviewed which implies that there is no data on the parametric structure of the process under study.

Keywords: priori information, H-processes, changing space dimension.

Introduction

A. Feldbaum [2] was one of the first scientists who paid attention to systems with various information and, accordingly, to the levels of a priori information. A.A. Feldbaum identified some levels of a priori information in his known book. They are as follows:

- systems with complete information;
- systems with maximum but incomplete information;
- systems with incomplete information;
- systems with active accumulation of information.

Later, Ya.Z. Tsypkin considered systems with parametric uncertainty. That is, it is a case when there is enough information a priori to determine the parametric structure of the object under investigation. He considers in [3] various problems in the theory of adaptive systems from these points of view. For example, an object model can be represented as a function of input actions.

$$x(t) = f(\vec{u}, t), \tag{1}$$

where x is the object's output, and $\vec{u}$ are input variables, t is time. If f(*) is known up to parameters, then (1) is as follows:

$$x(t) = f(\vec{u}, \vec{\alpha}, t), \tag{2}$$

where $\vec{\alpha}$ is a parameter vector.

This path is connected directly with approximations of the projection type. For us, the presence of an approximation of a local character is essential. A classic example of such an approximation of a function is the Lagrange interpolation polynomial

Denote a training sample (implementation of observations) as $(x_i, y_i), i = \overline{1, s}$. In this case, the function approximating (1), in particular, the regression function from observations, can be represented as follows:

$$x(t) = f(\vec{u}, t, \vec{x_s}, \vec{u_s}), \tag{3}$$

where $\vec{x_s}, \vec{u_s}$ are time vectors. $\vec{x_s} = (x_1, x_2, ..., x_s), \vec{u_s} = (u_1, u_2, ..., u_s)$.

As Y.Z. Tsypkin noted several times, a priori information is a tool for mathematical formulation of the problem, and current information is a tool for its solving.

We consider the levels of a priori information corresponding to the levels of parametric and non-parametric uncertainty in more detail.

Parametric uncertainty means the fact that a priori information is sufficient to determine reasonably the equation of the process under study with an accuracy of the parameter vector. The next steps in this operation are to estimate these parameters based on the current available information. This method is described in numerous monographs, in particular [3]. The application of the stochastic approximation method is described in some detail in the book written by Ya.Z. Tsypkin [3]. In particular, according to this book, algorithms of stochastic approximations are used to solve various problems of the adaptive systems theory [3].

Non-parametric uncertainty means the fact that a priori information is insufficient for a reasonable determination of the parametric equation of the process under investigation with an accuracy of the parameter vector when only its qualitative properties are known. For dynamic processes these features are linearity or non-linearity class; for inertialess systems with delay these features are unambiguous characteristics of different channels of an object or ambiguity. In this case, the objects can be described by the equations: $x(t) = f(\vec{u}, t, \vec{x_s}, \vec{u_s})$, or $F(\vec{x}(t), \vec{u}(t), t, \vec{x_s}, \vec{u_s}) = 0$, where f(.), F(.) are unknown.

In general, the investigated multidimensional system can be represented in Figure 1.

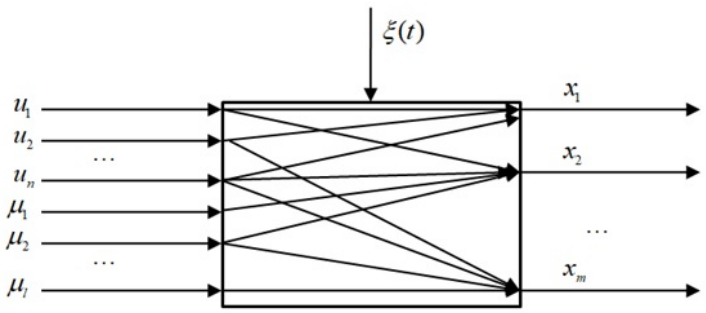


Figure 1: Multidimensional object

In the figure, the following designations are used: $\vec{u} = (u_1, \ldots, u_n)$ is n-dimensional vector of input controlled variables, $\vec{\mu} = (\mu_1, \ldots, \mu_l)$ l-dimensional vector of unmanaged input variables, but controlled, $\vec{x} = (x_1, \ldots, x_m)$ is m-dimensional vector of output variables.

Each component of the output variable vector can be represented using compound vectors.

Let's explain a term "a compound vector". It is a vector composed of some components of the corresponding input vectors. For example, the components of the vector of input variables $\nu(t)$(Figure 1) may have the form $\nu_t^1 = (u_t^1, \mu_t^3, u_t^2), \nu_t^2 = (u_t^2, u_t^4, \mu_t^1, \mu_t^3)$ and etc. The components of the compound vectors are directly dependent on the specifics of the process under investigation, the availability of a priori information about it, its characteristics, properties, etc.

The formation of compound vectors is carried out on the basis of the available a prior information. If it is absent, then the compound vector combines all the components of the corresponding variables.

A lot of really occurring processes in technologies and nature reveal the following feature: components of the vector of input and output variables turn out to be stochastically dependent (Figure 1).So, in particular, if the input variables of a multidimensional system are stochastically dependent, then the process under investigation takes place in a subregion of the "tube" structure. A frequent variant of such a structure is shown in Figure 2.

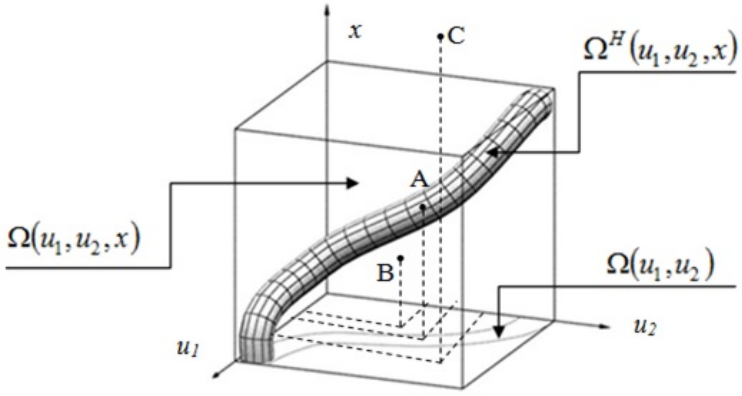


Figure 2: Process with a "tube" structure

In Figure 2, the following designations are used: $\Omega(u_1, u_2, x)$ is a region of the process without taking into account the stochastic dependence between the components of the vector of input variables; $\Omega^H(u_1, u_2, x)$ is a region of the process taking into account the stochastic dependence between the components of the vector of input variables.

Such processes are called H-processes.

It is natural to assume that the identifying process of an object in a parametric formulation should also be carried out taking into account the tube structure of the

object. In conclusion, we note that the nature H-processes are different from fractals and attractors.

If we interpret H-processes in a more general case as a function of several variables, then the variability of this function over time can be shown in the following chain of ratios operating in time:

$$\begin{cases} x = f(t, u_1, \mathbf{u_2}, \mathbf{u_3}, u_4, u_5) - T_1; \\ x = f(t, u_1, u_2, \mathbf{u_3}, u_4, u_5) - T_2; \\ x = f(t, u_1, u_2, u_3, u_4, u_5) - T_3; \\ x = f(t, u_1, u_2, u_3, u_4, u_5, u_6) - T_4; \\ x = f(t, u_1, u_2, u_3, u_4, u_5, u_6) - T_5; \\ x = f(t, u_1, u_2, u_3, u_4, u_5, u_6) - T_6; \\ x = f(t, u_1, \mathbf{u_2}, u_3, u_4, \mathbf{u_5}, u_6) - T_7; \\ x = f(t, \mathbf{u_1}, \mathbf{u_2}, u_3, u_4, \mathbf{u_5}, \mathbf{u_6}, u_7) - T_8. \end{cases} \tag{4}$$

Further the explanation of our designations is given. The darkest color (u_1) indicates the variables that have the strongest effect on x (possibly, functional dependence). The less obscure designation (u_1) indicates a weaker influence of the variable on x (perhaps, a stochastic dependence), and the variables u_1 and u_1 have a still weaker effect on x.

Thus, in actual H-processes, the influence of variable values changes significantly: some variables may lose their value, some variables may first lose value and then restore it, and some variables may appear for the first time, such as u_6, u_7.

The case when the components of the vector are stochastically related isimportant from the practical point of view. Then the dependency $x(\vec{u})$ can be described by a system of implicit functions of the following type:

$$F(\vec{x}, \vec{u}) = 0 \tag{5}$$

In this case, a system (4) can be represented as follows:

$$\begin{cases} F(t, x, u_1, \mathbf{u_2}, \mathbf{u_3}, u_4, u_5) = 0 - T_1; \\ F(t, x, u_1, u_2, \mathbf{u_3}, u_4, u_5) = 0 - T_2; \\ F(t, x, u_1, u_2, u_3, u_4, u_5) = 0 - T_3; \\ F(t, x, u_1, u_2, u_3, u_4, u_5, u_6) = 0 - T_4; \\ F(t, x, u_1, u_2, u_3, u_4, u_5, u_6) = 0 - T_5; \\ F(t, x, u_1, u_2, u_3, u_4, u_5, u_6) = 0 - T_6; \\ F(t, x, u_1, \mathbf{u_2}, u_3, u_4, \mathbf{u_5}, u_6) = 0 - T_7; \\ F(t, x, \mathbf{u_1}, \mathbf{u_2}, u_3, u_4, \mathbf{u_5}, \mathbf{u_6}, u_7) = 0 - T_8. \end{cases} \tag{6}$$

Let us examine Figure 1 from the point of view of the availability of a priori information, both parametric and non-parametric. Often a situation arises when, through some channels of a multidimensional object, its parametric structure can

be determined. And through other channels an investigator is under non-parametric uncertainty, that is, under conditions where the parametric structure cannot be determined due to lack of a priori information. Only some qualitative properties of the process under investigation are known. In this regard, a priori information is applied to both parametric and non-parametric levels of uncertainty. In such a situation, the development of the model is significantly different from the traditional approaches to the theory of identification.

Let us consider a grinding process of a specific product, a relatively typical for many industries. We mean grinding the clinker in three-chamber mills for dry grinding (the clinker is granules obtained as a result of burning raw mix, grinding which leads to the production of cement.

A dry grinding mill (Figure 3) is a cylindrical rotating drum, divided by grid partitions into three chambers loaded with grinding bodies: chamber I contains fairly large metal balls, chamber II contains smaller balls, chamber III contains cylinders (metal cylinders small size). The clinker entering the mill is crushed in chambers I, II, III and it is converted into cement. Thus, from a technological point of view, the entrance of the mill is clinker loading, and the output is cement.

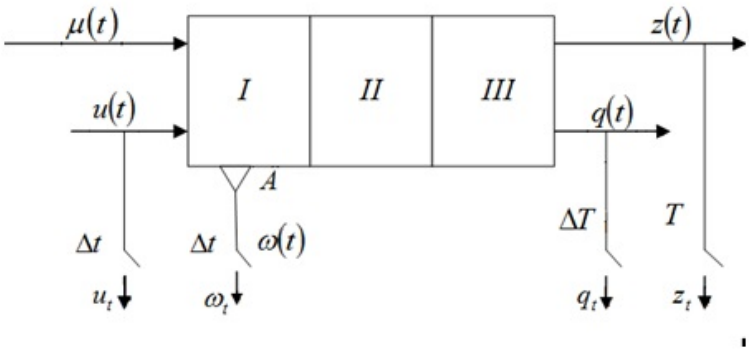


Figure 3: Scheme of the ball three-chamber mill

The following designations are introduced in Figure 3 $\mu(t)$ is uncontrolled input variable (grindability of the clinker); $u(t)$ is input variable controlled with a random error (load / number of the clinker); $\omega(t)$ is noise in the first chamber, monitored by induction sensor A in an interval of time Δt. It is used as an output signal of the grinding process in current control systems; q(t) is output of the mill (a technological parameter "fineness of grinding"), measured after an interval of time $\Delta T >> \Delta t$;$z(t)$ - the main indicator of the cement quality is the activity (strength of the cement beam under compression in 28 days after its mixing), monitored over a period of time $T >> \Delta T >> \Delta t$. The object time constant is about 5–7 minutes, u(t) and $\omega(t)$ in local analog control systems are monitored continuously, and it is discrete in digital control systems in an interval of time Δt (can be measured in a few seconds).

The control of the output variables $q(t), z(t)$ is carried out in the laboratory according to the technology regulated by the standard, with $\Delta T = 2$ hours and T = 28 days. Note that q(t) in this case is a technological indicator of the grinding

process, and $z(t)$ is the main indicator of the quality (grade) of cement (cement activity), which depends not only on the refinement of grinding q(t), but also on the performance of previous technological conversions: raw mix preparation, grinding, burning.

The variable $\mu(t)$ shown in Figure 3, is an important technological indicator of the grindability of the clinker, which significantly affects the grinding process of the latter, but cannot be measured.Of course, expert assessments of the grindability of clinker and the analysis of one of its granules by means of petrography and etc are possible. But everything requires much time. They are very laborious, unrepresentative, and they give rough average results. However, it was possible to develop a sufficiently high-quality system for controlling the process of dry grinding without measuring in this case such an indicator as the grindability of clinker. It was done as a result of the system analysis of the control unit and the technological unit of the grinding process, that led to a slightly different control system in the industry at that time.

Thus, the technology of variables control that affects significantly the grinding process is laborious in various time intervals, and the controlled process is a subject to various random factors, that leads to serious difficulties in modeling such processes. But this is a relatively simple technological process, which is typical for many industries, but even this process does not fit to the classical scheme of identification problem.

Hence one can understand the importance of a priori information for solving identification and control problems. Moreover, it can be seen that it can be different for different channels of the studied multidimensional system. Naturally, this leads to a special regard at the solution of the problems described above.

In conclusion, it should be noted that the levels of a priori information in the formulation of certain problems of identification and control inevitably lead to the necessity to use the relevant divisions of the control theory, which, of course, may be different.We have already talked about deterministic control, stochastic theories and etc. Each of the "floors" of a priori information corresponds to its own specific theory, both in the problems of the identification theory and in the problems of the control theory. The last level of non-parametric uncertainty should correspond to the theory of non-parametric systems [1]. The problems of identification and control under non-parametric uncertainty are also considered there.

References

[1] Feldbaum A. A. Fundamentals of the theory of optimal automatic systems. - M.: Fizmatgiz, 1963. - 553 p.

[2] Tsypkin Ya. Z. Adaptation and training in automatic systems. - M.: Science, 1968. - 399 p.

[3] Medvedev A.V. Fundamentals of the theory of non-parametric systems. - Krasnoyarsk: publisher - SibGAU, 2018. - 727 p.

Applying the method of moments to build the orthogonal series density estimator

Vladislav V. Branishti
Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russia
e-mail: branishti@sibsau.ru

Abstract

The task of nonparametric estimating the probability density function is considered in this paper. The author investigates the orthogonal series estimator when its coefficients are adjusted with the method of moments. It is given the method of adjusting the series length. Given estimator is compared with Rosenblatt – Parzen estimator.

Keywords: probability density function, nonparametric statistics, orthogonal series estimate, Rozenblatt – Parzen estimate, Monte Carlo method.

Introduction

The task of recovering the distribution of random variable, particularly, estimating the probability density function, is one of main objectives in mathematical statistics [5]. In this case the researcher often finds himself in a situation of nonparametric uncertainty [12]. The situation is complicated also by small amount of data. At present time there is range of nonparametric estimates of probability density, including kernel estimates, orthogonal series estimates, nearest neighbors method, etc. Due to this the task of choice the best method of estimation for given class of distributions and given amount of data n appears.

At present time the problem of comparison the quality of probability density function estimators is complex to compute even on model distributions. The quality of approximation, as a rule, is evaluated by a functional in a form of mathematical expectation of a random variable with unknown distribution. In this research the quality functional is numerically calculated by the Monte Carlo method.

1 Calculation the coefficients of orthogonal estimator with the method of moments

Let $x_1, \ldots, x_n$ be independent sample of continuous random variable ξ with unknown probability density $f(x)$ belonging to a Hilbert space $H(\Omega)$ of measurable real functions with the domain $\Omega \subseteq \mathbb{R}$. We choose in the space $H(\Omega)$ a complete orthonormal system (basis):

$$\{\varphi_0, \varphi_1, \ldots, \varphi_l, \ldots\}.$$

We call an *orthogonal series estimate* [3] of the probability density function the random function in form

$$\hat{f}(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \cdots + a_l\varphi_l(x),$$

where random variables l, a_0, ..., a_l are parameters of this estimate, that should be adjusted.

We will now estimate the coefficients a_0, ..., a_l, having in view that series length l is fixed. Assume that investigating random variable ξ has row moments of order $j = 1, \ldots, l$. Then values a_j we can find from equality between estimate of j-th raw moment, found with the orthogonal series estimate $\hat{f}(x)$, and j-th sample raw moment:

$$\int_{-\infty}^{+\infty} x^j \hat{f}(x)dx = \frac{1}{n}\sum_{i=1}^{n} x_i^j, \quad j = 0, \ldots, l. \tag{1}$$

The conditions (1) are equivalent to the system of linear equations about a_j:

$$\begin{cases} a_0(1,\varphi_0) + \cdots + a_l(1,\varphi_l) = 1 \\ a_0(x,\varphi_0) + \cdots + a_l(x,\varphi_l) = \hat{\nu}_1 \\ \vdots \\ a_0(x^l,\varphi_0) + \cdots + a_l(x^l,\varphi_l) = \hat{\nu}_l \end{cases}, \tag{2}$$

where $\hat{\nu}_j$ is j-th sample raw moment of the ξ:

$$\hat{\nu}_j = \frac{1}{n}\sum_{i=1}^{n} x_i^j,$$

(f, g) is inner product in the space $H(\Omega)$:

$$(f,g) = \int_{\Omega} f(x)g(x)dx.$$

Using matrix notations

$$\mathbf{B} = \begin{pmatrix} (1,\varphi_0) & \ldots & (1,\varphi_l) \\ \vdots & \ddots & \vdots \\ (x^l,\varphi_0) & \ldots & (x^l,\varphi_l) \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} a_0 \\ \vdots \\ a_l \end{pmatrix}, \quad \hat{\boldsymbol{\nu}} = \begin{pmatrix} \hat{\nu}_0 \\ \vdots \\ \hat{\nu}_l \end{pmatrix},$$

we can rewrite the system (2) by such a way:

$$\mathbf{Ba} = \hat{\boldsymbol{\nu}}.$$

Now, if the determinant $|\mathbf{B}| \neq 0$ then there exists unique solution

$$\mathbf{a} = \mathbf{B}^{-1}\hat{\boldsymbol{\nu}}. \tag{3}$$

We suggest to use the elements of found column matrix $\mathbf{a}$ as estimates of coefficients for $\hat{f}(x)$.

2 Calculation of series length

The choice of series length l significantly affects the quality of orthogonal series estimate. We will characterize the quality of estimator $\hat{f}(x)$ by mean global quadratic approximation error [6], which we will minimize by l:

$$Q\{\hat{f}\} = M\left\{\left\|\hat{f} - f\right\|^2_{H(\Omega)}\right\} \to \min_l.$$

After simplifying this functional has a form:

$$Q\{\hat{f}\} = M\left\{\|\hat{f}\|^2 - 2(\hat{f}, f)\right\} + \|f\|^2.$$

The last summand in received expression does not depend on l, so it can be omit within minimization. Thus, for adjusting the series length we define the functional

$$W\{\hat{f}\} = M\left\{\|\hat{f}\|^2 - 2(\hat{f}, f)\right\}. \tag{4}$$

As shown in [10], the minimum of functional (4) is achieved with finite l. When we use the formula (3) for estimation the coefficients a_j, the functional (4) takes the form:

$$W\{\hat{f}\} = \mathrm{tr}\mathbf{B}^{-1}\left(M\left\{\hat{\boldsymbol{\nu}}\hat{\boldsymbol{\nu}}^T\right\}\left(\mathbf{B}^{-1}\right)^T - 2\boldsymbol{\nu}\boldsymbol{\alpha}^T\right), \tag{5}$$

where

$$\boldsymbol{\nu} = (\nu_0, \ldots, \nu_l)^T, \quad \boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_l)^T,$$

$\nu_j = (x^j, f) = \int_{-\infty}^{+\infty} x^j f(x)dx$ – j-th raw moment of random variable ξ;

$\alpha_j = (\varphi_j, f) = \int_{-\infty}^{+\infty} x^j f(x)dx$ – j-th Fourier coefficient of series of the density $f(x)$ by basis $\{\varphi_j\}$.

During research the unbiased sample estimate of the functional (5) was built:

$$\hat{W}_l = \mathrm{tr}\mathbf{B}^{-1}\left(\hat{\boldsymbol{\nu}}\hat{\boldsymbol{\nu}}^T\left(\mathbf{B}^{-1}\right)^T - 2\hat{\mathbf{G}}\right), \tag{6}$$

where

$$\hat{\mathbf{G}} = \begin{pmatrix} g_{0,0} & \cdots & g_{0,l} \\ \vdots & \ddots & \vdots \\ g_{l,0} & \cdots & g_{l,l} \end{pmatrix},$$

$$g_{j_1,j_2} = \frac{1}{n(n-1)}\left(\sum_{i=1}^{n} x_i^{j_1}\sum_{i=1}^{n}\varphi_{j_2}(x_i) - \sum_{i=1}^{n} x_i^{j_1}\varphi_{j_2}(x_i)\right).$$

Then we can choose series length estimate $\hat{l}$ by minimization the (6):

$$\hat{l} = \arg\min_l \hat{W}_l.$$

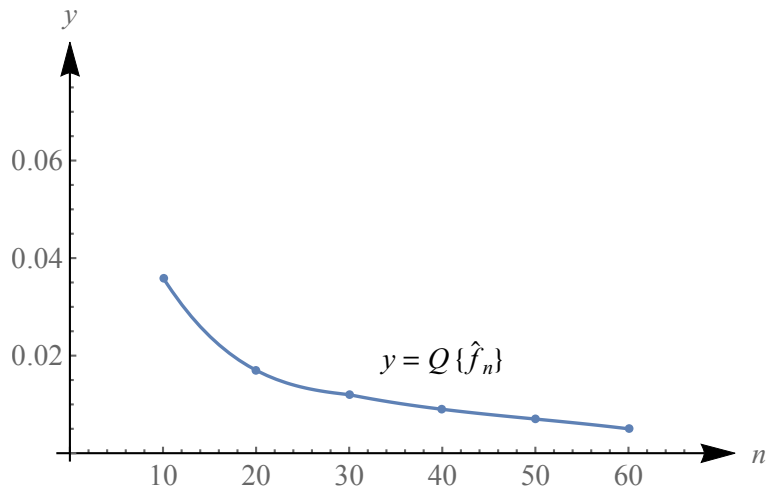


Figure 1: Dependence of estimator quality on sample amount

Due to asymptotic features of the method of moments [4], built estimator of probability density function is consistent, and it is also confirmed in numerical experiments (fig. 1).

Fig. 1 shows the dependence of quality functional (2) on sample amount n withing the task of recovering the standard normal distribution, where $\{\varphi_j(x)\}$ is the Hermite basis [5, ch. 12]. Values of the functional (2) was calculated by Monte Carlo method described in [1].

Also during research the quality comparison between suggested estimator and kernel Rosenblatt – Parzen estimator [11] has been implemented. The spread parameter of the kernel estimator is calculated by method of minimizing the estimate of quality functional [1] or maximum likelihood method [7] (fig. 2).

On fig. 2 we made following notations:

- $\hat{f}_n^{(1)}(x)$ is suggested orthogonal series estimator;
- $\hat{f}_n^{(2)}(x)$ is Rozenblatt – Parzen estimator in which

$$\hat{h} = \arg\min_h \left(\frac{1}{hn^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \tau\left(\frac{x_i - x_j}{h}\right) - \frac{2}{hn(n-1)} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \Phi\left(\frac{x_j - x_i}{h}\right) \right),$$

$$\tau(x) = \int_{\mathbb{R}} \Phi(z)\Phi(z+x)dz,$$

$\Phi(z)$ – Gaussian kernel function;

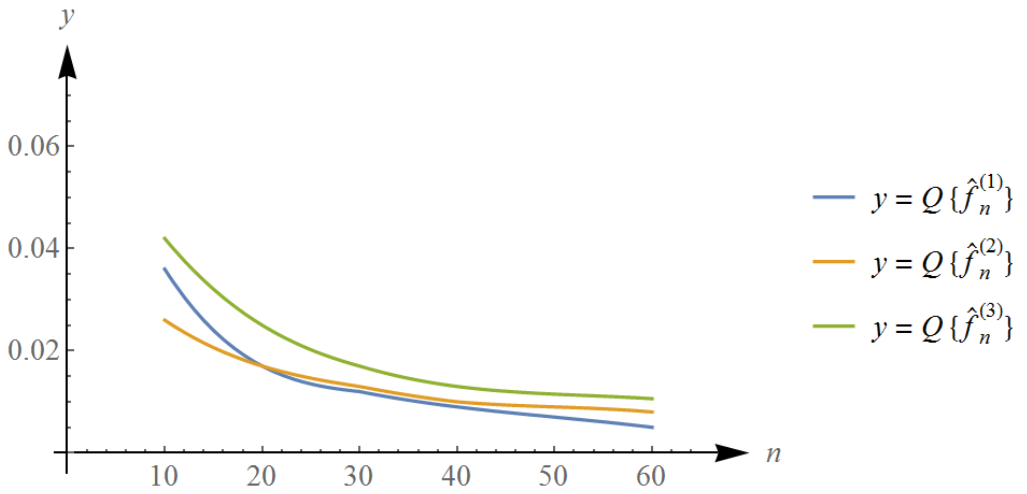


Figure 2: Comparison the quality of density estimators

- $\hat{f}_n^{(3)}(x)$ is Rozenblatt – Parzen estimator in which

$$\hat{h} = \arg\max_h \sum_{i=1}^{n} \ln \sum_{\substack{j=1 \\ j\neq i}}^{n} \Phi\left(\frac{x_i - x_j}{h}\right) - n\ln((n-1)h).$$

This comparison is implemented to recover the probability density of standard normal distribution. As we can see, suggested estimator gives lesser mean global square error approximation than Rozenblatt – Parzen estimator in sufficiently wide range of sample amount n (at least when $20 \leq n \leq 60$).

Conclusions

Thus, in the paper we have considered applying the method of moments to build the orthogonal series estimator of probability density function of continuous random variable. It should be noted that to use the orthogonal series estimator requirement that true density belongs to Hilbert space is not too restrictive, because, as shown in [2], for each probability density function $f(x)$ exists a Hilbert space in form $L_{2,w}(\Omega)$ that contains $f(x)$.

Given method of comparison nonparametric estimators can be also applied to compare other nonparametric estimators, e.g. histogram, k-nearest neighbors [9], etc.

References

[1] Branishti V. V. (2014). O parametricheskom otsenivanii funktsii plotnosti veroyatnosti [On parametric estimation of probability density function]. *Nauchno-*

tekhnicheskiy vestnik Povolzhya. No. **1**, pp. 13–16.

[2] Branishti V. V. (2017). On some Properties of Weighted Hilbert Spaces. *Journal of Siberian Federal University. Mathematics & Physics*. Vol. **10** (4), pp. 410–421.

[3] Chentsov N. N. (1962). Otsenka neizvestnoi plotnosti raspredeleniya po nablyudeniyam [Estimate of unknown probability density on observations]. *Dokl. AN SSSR*. Vol. **147**, no. 1, pp. 45–48.

[4] Cramér H. (1946). *Mathematical methods of statistics*. Princeton University Press, Princeton.

[5] Devroye L., Gyorfi L. (1985). *Nonparametric density estimation. The L_1 view*. John Wiley & Sons, New York.

[6] Epanečnikov V. A. (1969). Nonparametric estimation of a multidimensional probability density. *Teor. Veroyatnost. i Primenen*. Vol. **14**, issue 1, pp. 156–161.

[7] Lapko A. V., Lapko V. A. (2010). *Neparametricheskiye modeli i algoritmy obrabotki informatsii* [*Nonparametric models and algorithms of information processing*]. SibGAU Publ., Krasnoyarsk.

[8] Lapko A. V., Medvedev A. V., Tishina Ye. A. (1975). K optimizatsii nekotorykh neparametricheskikh otsenok [On optimization of some nonparametric estimates]. *Primeneniye vychislitel'nykh mashin v sistemakh upravleniya nepreryvnym proizvodstvom*. Frunze, pp. 93–107.

[9] Loftsgaarden D. O., Quesenberry C. P. (1965). A Nonparametric Estimate of a Multivariate Density Function. *Ann. Math. Statist*. Vol. **36**, no. 3, pp. 1049–1051.

[10] Novosyolov A. A. (1979). *Ob optimal'nom vybore struktury functsii plotnosti veroyatnosti i regressii* [*On optimal choice of structure of probability density function and regression*]. VC SO RAN, Krasnoyarsk.

[11] Parzen E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist*. Vol. **35**, no. 3, pp. 1065–1076.

[12] Tarasenko F. P. (1976). *Neparametricheskaya statistika* [*Nonparametric statistics*]. Tomsk University Publ., Tomsk.

Robust correlation coefficients based on weighted maximum likelihood method

Oleg S. Cherepanov
Kurgan State Universirt, Kurgan, Russian Federation
e-mail: ocherepanov@inbox.ru

Abstract

In the work new robust correlation coefficients on the basis of weighted maximum likelihood method are suggested. The research is conducted on the efficiency of proposed estimates on the class of elliptical distribution. It is shown that the estimates got have a high level of efficiency while having outliers of Tukey's model.

Keywords: correlation, robust estimates, weighted maximum likelihood method, adaptive estimates.

Introduction

The task of the estimation of parameters of statistical models is one of the main tasks of mathematical statistics. By now quite a number of estimates, having different qualities, have been proposed. The choice of the estimates depends on the information given a priori. The researches show that in real observations outliers may often present (abnormal observations), which can degrade classical estimates. If there is such information, we deal with robust statistics. Within this theory, estimates resistant to outliers on the basis of different criteria have been suggested [1]-[3]. In robust statistics much attention is paid to the construction of robust estimates of distribution parameters and a little less attention is paid to the estimation of links between random variables. It is known that sample Pearson correlation coefficient is resistant to the outlier presence and its proposed robust analogues [1]-[6]. do not have high efficiency in the conditions of outlier absence. The criteria of efficiency and resistance appear to be competitive. Often effective estimates own low robust qualities and, vice a versa, high robust estimates own low efficiency without outliers. So, it is necessary to get "compromise" estimates, capable to adjust to the form and level of "contamination" of observations, staying robust to outliers and at the same time having high efficiency.

1 Estimates of weighed maximum likelihood method

Let's consider two-dimensional random variable (X, Y), having elliptical distribution [7] with joint density of the form:

$$f(x,y,\theta) = \frac{c}{s_x s_y \sqrt{1-p^2}}\, g\left(\frac{1}{1-p^2}\left[\left(\frac{x-\mu_x}{s_x}\right)^2 - 2p\left(\frac{x-\mu_x}{s_x}\right)\left(\frac{y-\mu_y}{s_y}\right) + \left(\frac{y-\mu_y}{s_y}\right)^2\right]\right), \qquad (1)$$

where $g(x)$ is a generating function ($R \to R^+$), fulfilling condition $\int_0^\infty g(x)\,dx < \infty$;
$\theta = (\mu_x, \mu_y, s_x, s_y, p)^T$,
μ_x and μ_y location parameters of random variables X, Y respectively;
s_x and s_y are scale parameters of random variable X, Y respectively;
p is a correlation coefficient;
c is a norming quantity.

According to the weighed maximum-likelihood method (WMLM) [8] the parameter estimate θ of a random variable X with density $f(x, \theta)$ will be determined by equation of the form:

$$\sum_{i=1}^{N} \left(\frac{\partial log\ f(x_i, \theta)}{\partial \theta} + \beta_\theta \right) f^l(x_i, \theta) = 0, \tag{2}$$

where β_θ, is a parameter, answering for unbiasedness of the estimate,
l is a radical parameter, $l \in [0; 1]$.

Using the weighed maximum likelihood method (2) for parameter estimation of two-dimensional random variable (X, Y) with density (1), defining parameters β, as well as using identity substitution, the estimates of the following from will be got:

$$\begin{aligned}
&\sum_{i=1}^{N} (x_i - \hat{\mu}_x) w(x_i, y_i, \hat{\theta}) f^l(x_i, y_i, \hat{\theta}) = 0,\\
&\sum_{i=1}^{N} (y_i - \hat{\mu}_y) w(x_i, y_i, \hat{\theta}) f^l(x_i, y_i, \hat{\theta}) = 0,\\
&\sum_{i=1}^{N} \left[\left(\frac{x_i - \hat{\mu}_x}{s_x} \right)^2 w(x_i, y_i, \hat{\theta}) + \frac{1}{l+1} \right] f^l(x_i, y_i, \hat{\theta}) = 0,\\
&\sum_{i=1}^{N} \left[\left(\frac{y_i - \hat{\mu}_y}{s_y} \right)^2 w(x_i, y_i, \hat{\theta}) + \frac{1}{l+1} \right] f^l(x_i, y_i, \hat{\theta}) = 0,\\
&\sum_{i=1}^{N} \left[\left(\frac{x_i - \hat{\mu}_x}{s_x} \right) \left(\frac{y_i - \hat{\mu}_y}{s_y} \right) w(x_i, y_i, \hat{\theta}) + \frac{p}{l+1} \right] f^l(x_i, y_i, \hat{\theta}) = 0,\\
&w(x, y, \theta) = \frac{2f'(x, y, \theta)}{f(x, y, \theta)},
\end{aligned} \tag{3}$$

$$f(x, y, \theta) = \frac{c}{s_x s_y \sqrt{1 - p^2}}\, g' \left(\frac{1}{1 - p^2} \left[\left(\frac{x - \mu_x}{s_x} \right)^2 - 2p \left(\frac{x - \mu_x}{s_x} \right) \left(\frac{y - \mu_y}{s_y} \right) + \left(\frac{y - \mu_y}{s_y} \right)^2 \right] \right).$$

If for joint density of the distribution $f(x, y, \theta)$ marginal densities of the distribution can be received analytically, the system of equation (3) may be divided into two systems of equation for the parameter estimation of random variables X and Y and non-linear equation for estimation of correlation coefficient.

The efficiency of estimates (3) depends on the radical parameter l. At $l = 0$ we get estimates of maximum likelihood method, at $l = 0.5$ - radical estimates [3] and at $l = 1$ – estimates of maximum stability [3]. The radical parameter defines robust qualities and efficiency of estimates. The higher the radical parameter is, the higher the robust qualities of the estimate are, though in the conditions of outlier absence the efficiency of such estimates, as the researches show [9]-[12], is not high. Under the conditions of coincidence of a priori distribution with real distribution the radical parameter must approach to 0 (estimate of maximum likelihood method). However estimates of maximum likelihood method of distribution parameters with light tail while having outliers may own an extremely low efficiency. That's why a procedure of estimate adaptation of WMLM is required on the radical parameter to the form and a fraction of outliers.

Let only correlation coefficient p be estimated, while all the other parameters are parasitic. In this case the optimal value of the radical parameter will be defined through the solution of optimization problem of the form:

$$l^* = \min_{l \in [0;1]} V(\hat{p}, l),$$

where $V(\hat{p}, l)$ is a mean square error of estimate $\hat{p}$ from true value.
In practice when there is only one sample it is suggested to use its estimation on the basis of bootstrap method instead of mean square error of estimate [13].

2 Study of estimate efficiency

2.1 Two-dimensional elliptical distributions

In the work the research is carried out on the efficiency of correlation coefficient p estimates for distributions with different degree of tail stretching, belonging to the class of elliptical distributions: two-dimensional 4th degree generalized normal distribution (GND4), two-dimensional normal distribution (ND), two-dimensional Laplace distribution (LD), two-dimensional Cauchy distribution (CD). Analytic expressions of generic functions and norming quantity values are given in table 1.

Table 1: Elliptical two-dimensional distributions

Distribution	Generic function	Norming quantity
4th degree GND	e^{-x^2}	$\frac{2}{\pi\Gamma(0.5)}$
Normal distribution	$e^{-x/2}$	
Laplace distribution	$e^{-\sqrt{x}}$	$\frac{1}{2\pi}$
Cauchy distribution	$(1+x)^{3/2}$	

Graphic representations of joint densities of distributions are given in figures 1-4.

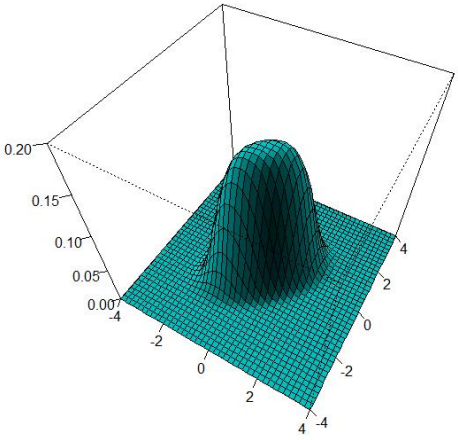


Figure 1: Density of two-dimensional 4th generalized normal distribution

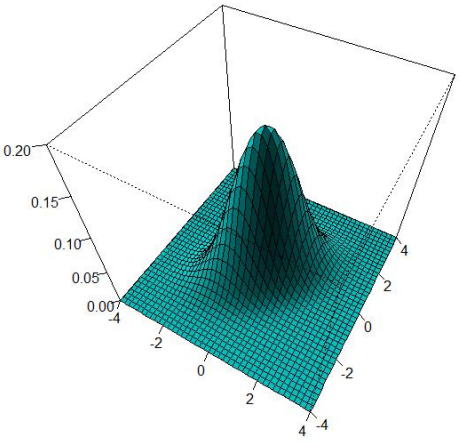


Figure 2: Density of two-dimensional normal distribution

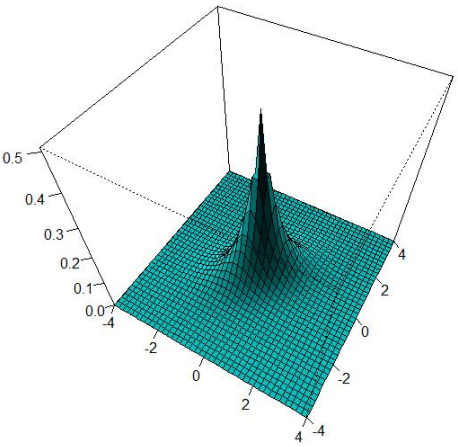


Figure 3: Density of two-dimensional Laplace distribution

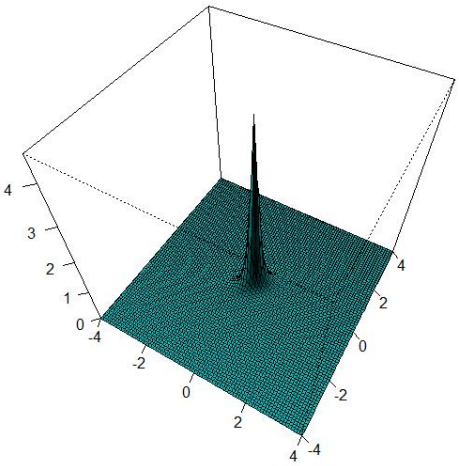


Figure 4: Density of two-dimensional Cauchy distribution

Within the research of the received estimates scale parameters of the studied distributions have been estimated so that quantiles of level 0.95 of marginal distributions coincided.

2.2 Models of outliers

To research the behavior of estimates while having outliers Tukey's supermodel has been used [1], [2], [4] with a fraction of outliers ε:

1. Model of asymmetrical outliers:

$$f(x,y) = (1-\varepsilon)g(x,y,\mu_x,\mu_y,s_x,s_y,p) + \varepsilon g(x,y,\mu_x-8,\mu_y-8,s_x,s_y,p).$$

2. Model of symmetrical outliers:

$$f(x,y) = (1-\varepsilon)g(x,y,\mu_x,\mu_y,s_x,s_y,p) + \varepsilon g(x,y,\mu_x,\mu_y,3s_x,3s_y,p).$$

3. Model of outliers on parameter p:

$$f(x,y) = (1-\varepsilon)g(x,y,\mu_x,\mu_y,s_x,s_y,p) + \varepsilon g(x,y,\mu_x,\mu_y,s_x,s_y,-p).$$

Graphic representations of samples of mixture of two-dimensional 4th degree generalized normal distributions are given in figures 5-7.

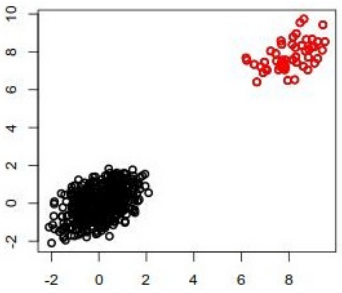


Figure 5: Model of asymmetrical outliers

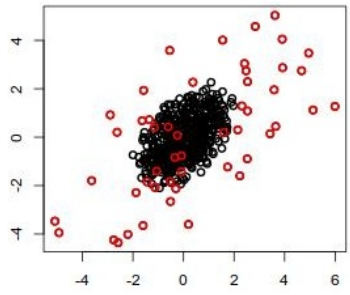


Figure 6: Model of symmetrical outliers

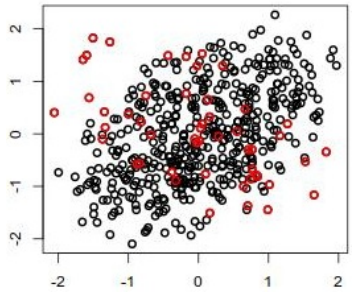


Figure 7: Model of outliers on parameter p

2.3 Results of research of estimate efficiency

The efficiency of received estimates is compared to the efficiency of classical estimates. Let us introduce the following notation:

AED4 – estimation of parameter p of two-dimensional 4th degree generalized normal distribution on the basis of the weighed maximum likelihood method with adaptation on the radical parameter.

AEND – estimation of parameter p of two-dimensional normal distribution on the basis of the measured maximum-likelihood method with adaptation on the radical parameter.

AELD – estimation of parameter p of two-dimensional Laplace distribution on the basis of the measured maximum-likelihood method with adaptation on the radical parameter.

AECD – estimation of parameter p of two-dimensional Cauchy distribution on the basis of the measured maximum-likelihood method with adaptation on the radical parameter.

SCP – sample Pearson correlation coefficient.

FQ – quick robust estimate of Shevlyakov-Smirnov [5], [6].

It is important to note that while modeling package robcor was used [14], in which FQ-estimate is realized. Parameters for FQ- estimate were chosen by default.

In tables 3-6 a part of results of research of estimate efficiency of correlation coefficient on different samples at the presence and absence of outliers of Tukey's

model is presented. Parameters of real distributions are given in table 2. Sample scope has comprised 1000 observations, the number of samples is 500.

Relative estimate efficiency has been defined within the chosen estimates:

$$\epsilon(\hat{p}, F) = \frac{min(MSE(\hat{p}_i, F))}{MSE(\hat{p}, F)},$$

where $MSE(\hat{p}, F)$ is a mean-square error of estimate $\hat{p}$ on distribution F.

The reference was considered the estimate, which has a minimum MSE among the studied estimates.

Table 2: Parameters of real distributions

Parameter	μ_x	μ_y	s_x	s_y	p	ε
4th degree generalized normal distribution	0	0	1.6	1.6	0.5	0.1
Normal distribution	0	0	1	1	0.5	0.1
Laplace distribution	0	0	0.6	0.6	0.5	0.1
Cauchy distribution	0	0	0.2	0.2	0.5	0.1

Table 3: Efficiencies and mean square error of estimates at two-dimensionalc 4th degree generalized normal distribution and presence of asymmetrical outliers

Estimates	**AED4**	AEND	AELD	AECD	SCP	FQ
MSE	**0.000422**	0.000593	0.001243	0.001719	0.197704	0.024029
Efficiency	**1.00**	0.71	0.34	0.25	0.00	0.02

Table 4: Efficiencies and mean square error of estimates at two-dimensional normal distribution

Estimates	AED4	**AEND**	AELD	AECD	**SCP**	FQ
MSE	0.000589	**0.000520**	0.000589	0.001062	**0.000520**	0.000642
Efficiency	0.88	**1.00**	0.88	0.49	**1.00**	0.81

Conclusions

According to the results of the researches it is possible to make the following conclusions:

Table 5: Efficiencies and mean square error of estimates at two-dimensional Laplace distribution and presence of symmetrical outliers

Estimates	AED4	AEND	**AELD**	AECD	SCP	FQ
MSE	0.002341	0.001700	**0.001333**	0.002117	0.009875	0.002860
Efficiency	0.57	0.78	**1.00**	0.63	0.14	0.47

Table 6: Efficiencies and mean square error of estimates at two-dimensional Cauchy distribution and presence of outliers on parameter p

Estimates	AED4	AEND	AELD	**AECD**	SCP	FQ
MSE	0.009596	0.008149	0.007794	**0.007391**	0.130750	0.008602
Efficiency	0.77	0.91	0.95	**1.00**	0.05	0.86

1. Sample Pearson correlation coefficient has an extremely low efficiency at the presence of outliers.
2. Quick robust estimate of Shevlyakov-Smirnov has quite a high efficiency on all the models, excepting for the model with asymmetrical outliers. The doubtless advantage of this estimate is its low level of calculation in relation to WMLM.
3. The estimates of the weighed maximum likelihood method have a high efficiency only on condition of coincidence of a priori distribution with basic real distribution. Otherwise estimate of the weighed maximum likelihood method may have not a very high efficiency even with the use of adaptation procedure. So, it is necessary to use adaptation not only to outliers but to the form of basic distribution of Tukey's model with the use of non-parametric estimates of density [12].
4. The estimates of WMLM are not able to suppress inner outliers (outliers on parameter p) because of "compressed" qualities of weight function $f^l(x_i, \theta)$.

So, the received robust estimates of WMLM may be used effectively for estimation of correlation coefficient in the conditions of semi-parametric level of a priori information at the presence or absence of symmetrical and asymmetrical outliers. The proposed procedure to define optimal value of radical parameter contributes to the adaptation of the estimates of WMLM to a fraction of outliers, increasing their efficiency.

References

[1] Hampel F., Ronchetti E., Rousseeuw P., Shtahel W. (1989). *Robust statistics. Approach on the basis of influence function*. Mir, Moscow.

[2] Huber P. (1984). *Robust statistics*. Mir, Moscow.

[3] Shurygin A.M. (2000). *Applied statistics. Robustness. Estimation. Forecast*. Finance and statistics, Moscow.

[4] Shulenin V.P. (2012). *Mathematical statistics. Part 3*. Publishing house TSU, Tomsk.

[5] Shevlyakov G.L., Smirnov P.O. (2010). Highly efficient robust estimators of a correlation coefficient for bivariate independent component distributions *Book of Abstracts: International Conference on Robust Statistics (ICORS 2010)*, pp. 93–94.

[6] Shevlyakov G.L., Smirnov P.O. (2011). Robust estimation of the correlation coefficient: An attempt of survey *Austrian Journal of Statistics*. Vol. **40**, pp. 147–156.

[7] Frahm G. (2004). *Generalized elliptical distributions: Theory and Applications. Ph.D.*. Germany, Cologne.

[8] Simakhin V. A. (2008). The measured maximum-likelihood method *Cybernetics and high technologies of the XXI century: materials of the IX international scientific technical conference*. Vol. **2**, pp. 661-672.

[9] Batrakov P. A., Cherepanov O.S. (2014). The research of estimates of scale parameter of the measured maximum-likelihood method. *Vestnik of Omsk State University. Management, computing and information technology.*. Vol. **2**, pp. 18–22.

[10] Simakhin V. A. (2013). Adaptive estimates of location parameter *Vestnik of Omsk state university. Management, computing and information technology*. Vol. **22**, pp. 131–137.

[11] Simakhin V.A., Cherepanov O.S. (2011). The research of estimates of the measured maximum-likelihood method *Vestnik of Kurgan state university. Series: Technical sciences*. Vol. **6**, pp. 72-77.

[12] Cherepanov O.S. (2016). *Robust parameter estimates on the basis of the measured maximum-likelihood method:thesis of candidate of physical and mathematical sciences*. Russian Federation, Tomsk.

[13] Davison A. C., Hinkley D. V. (1997). *Bootstrap Methods and Their Application*. Great Britain, Cambridge.

[14] Smirnov, P. O. (2013). *Robust correlations. R package version 0.1-5*. Austria. Vienna.

About non-parametric algorithms identification of inertialess systems

Svetlana S. Andoni, Vladislav V. Andoni, Anastasia V. Shishkina and Daria I. Yareschenko
Siberian Federal University, Krasnoyarsk, Russia
e-mail: chsvetlanas@gmail.com

Abstract

Regression recovery from observations with errors can be performed within the parametric or nonparametric setting. The problems of nonparametric estimation of regression characteristics by measurements of variables with errors under some features are considered below.

Keywords: : mutually ambiguous characteristics, nonparametric model, nonparametric estimates of the regression function, robust estimation.

Introduction

The identification tasks [1,2] of inertialess systems with delay, as mentioned above, are often close to estimating regression characteristics according to observation of input-output variables. These tasks are focused on the restoration of the regression function according to the observations, when the studied processes are described by mutually ambiguous characteristics [3-6]. These tasks are reduced to the problem of approximation, the feature of which is the absence of a priori information about the parametric structure of the model of the process under study. A non-parametric evaluation of mutually ambiguous characteristics [7, 8] and a robust estimation technique [9–12] are proposed.

1 Nonparametric estimates of mutually ambiguous regression function from observations

To restore the regression function, the non-parametric Nadaraya-Watson estimate is used, for a one-dimensional case, as follows (1):

$$Y_s(x) = \frac{\sum_{i=1}^{s} y_i \Phi(\frac{x-x_i}{c_s})}{\sum_{i=1}^{s} \Phi(\frac{x-x_i}{c_s})}, \tag{1}$$

where $\Phi(v)$ - the core - is a finite bell-shaped function satisfying the properties:

$$0 < \Phi(v) < \infty \ \forall v \in \bigcap(v),\ \frac{1}{c_s}\int \Phi\left(\frac{x-x_i}{c_s}\right) dx = 1, \lim_{n\to\infty} \frac{1}{c_s}\Phi\left(\frac{x-x_i}{c_s}\right) = \delta\left(x - x_i\right),$$

c_s- blur parameter with properties:

$$c_s > 0, \lim_{s\to\infty} s(c_s)^k = \infty, \lim_{s\to\infty} c_s = 0$$

if x is a k-dimensional vector, then the formula takes the form (2):

$$Y_s(x) = \frac{\sum_{i=1}^{s} y_i \prod_{j=1}^{k} \Phi(\frac{x_j - x_j^i}{c_s})}{\sum_{i=1}^{s} \prod_{j=1}^{k} \Phi(\frac{x_j - x_j^i}{c_s})}. \quad (2)$$

When restoring the mutually ambiguous regression function, the Nadaraya-Watson estimate should be changed as follows [3, 14]:

$$Y_s(x_t) = \frac{\sum_{i=1}^{s} y_i \Phi(\frac{x_t - x_i}{c_s}) \Phi(\frac{x_{t-1} - x_{i-1}}{c_s}) \Phi(\frac{y_{t-1} - y_{i-1}}{c_s})}{\sum_{i=1}^{s} \Phi(\frac{x_t - x_i}{c_s}) \Phi(\frac{x_{t-1} - x_{i-1}}{c_s}) \Phi(\frac{y_{t-1} - y_{i-1}}{c_s})}, \quad (3)$$

where x_{t-1}, y_{t-1} are the coordinates of the regression function at the previous step of its estimation.

It is advisable (3) to correct, as shown by numerous computational experiments, as follows (4):

$$Y_s(x_t) = \frac{\sum_{i=1}^{s} y_i \Phi(\frac{x_t - x_i}{c_s}) \Phi^0(\frac{x_{t-1} - x_{i-1}}{c_s}) \Phi^0(\frac{y_{t-1} - y_{i-1}}{c_s})}{\sum_{i=1}^{s} \Phi(\frac{x_t - x_i}{c_s}) \Phi^0(\frac{x_{t-1} - x_{i-1}}{c_s}) \Phi^0(\frac{y_{t-1} - y_{i-1}}{c_s})}, \quad (4)$$

where $\Phi^0(v)$ repeats with accuracy $\Phi(v)$, and $\Phi^0(v) = 1$, if $v < 1$ and 0 in other cases. In this case, $\Phi^0(v)$ will not affect the recovery error, but will allow "fixing" the previous data area at the estimated point as it moves along the chosen trajectory.

2 Computational Experiments

Below we present the results of some computational experiments for reconstructing mutually ambiguous characteristics from observations. When conducting a computational experiment, mutually ambiguous characteristics may have different shapes: circles, ellipses, hysteresis loops, Cassini oval, cardioid, and other curves.

We considered cases with the addition of random noise h on the values of y in computational experiments. The interference to each observation of the output variable forms as follows (5):

$$h_i = l y_i \xi, \quad (5)$$

where $\xi \in [-1, 1]$, level of interference $l = 0\%; 5\%; 10\%$.

As a non-parametric evaluation accuracy criterion, we used the ratio (6):

$$w = \sum_{i=1}^{s} |y_i - y_s(x_i)| \Big/ \sum_{i=1}^{s} |y_i - \overline{y}|, \quad (6)$$

where $\overline{y} = \frac{1}{s}\sum_{i=1}^{s} y_i$ - average; $y_s(x_i)$ - nonparametric evaluation; y_i - the real sample.

We present the results of a numerical study illustrating the efficiency of the algorithm. A triangular core was used as the bell-shaped function. The algorithm was tested on training samples of various sizes.

Let us designate in all figures by the digit (1) - the training sample, (2) - non-parametric estimation.

Figures 1 and 2 show the operation of the algorithm (3). Figures 3 and 4 show the operation of the algorithm (4).

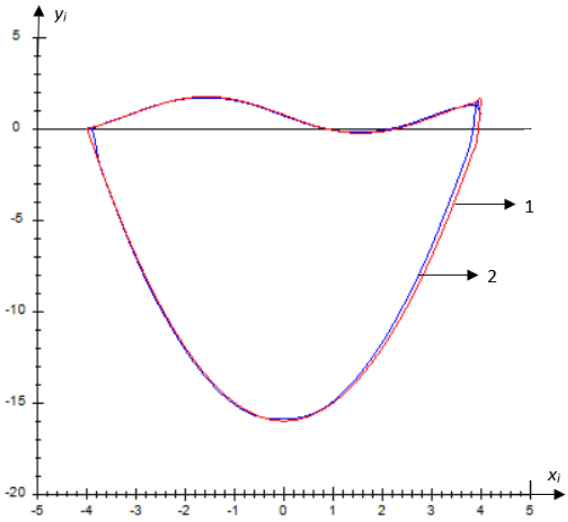


Figure 1: S=100; w=0,0205

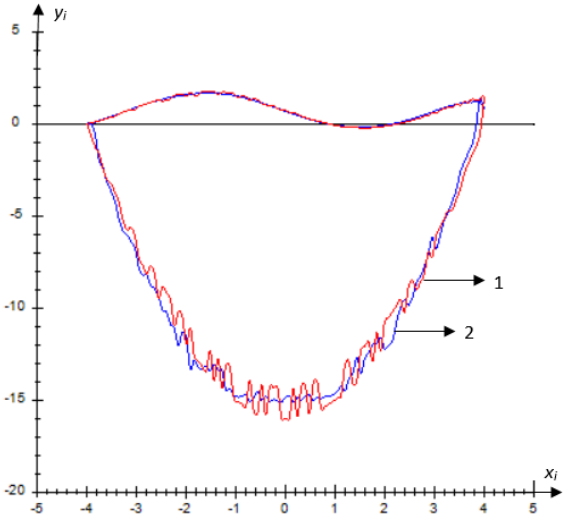


Figure 2: S=100; l=10%; w=0,0404

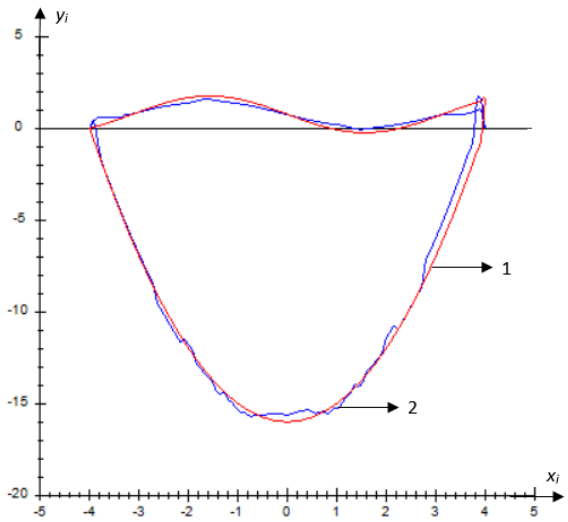


Figure 3: S=100; l=0%; w=0,0062

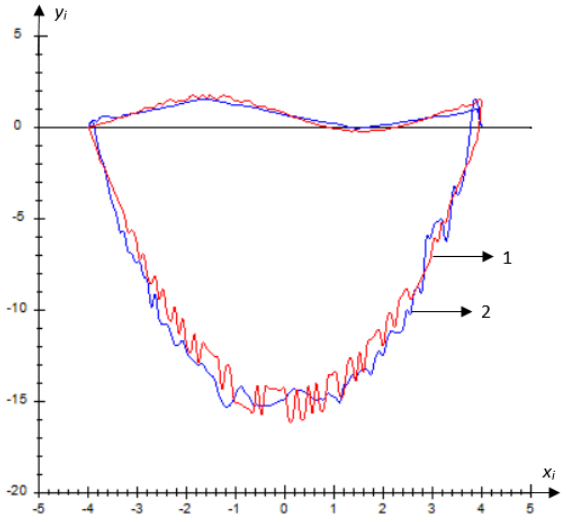


Figure 4: S=100; l=10%; w=0,0098

These figures show some fragments of numerous computational experiments. Algorithms (3) and (4) were used. They differ from the well-known non-parametric

algorithms for estimating the Nadaraya-Watson regression function by observation. Some modifications of non-parametric estimates are given, in such conditions attention is paid to the method of bypassing the entered non-parametric estimates along the trajectory determined by the elements of the training set.

In other words, algorithms suitable for recovering ambiguous dependencies that are described by more complex curves are proposed. We only know: a sample of observations of the process under study. Such identification tasks may appear when building robotic devices when driving in unknown terrain.

3 Non-parametric robust estimation

When solving the problem of identification with inertia-free objects with a delay, sometimes there are overshoots that affect the variable characteristics of the object. Actually, there is a robust statistics, which designed to reduce the impact of a miss on the further evaluation of statistical characteristics as subsequent measurements arrive.

Below are a few approaches, different from [11,12], which restore the values of the estimates of various statistical characteristics. The impact of emissions will disappear as relevant variable measurements become available. The differences in the considered approach consist in the detection and exclusion of an overshoot from the training sample [15], which is used for the statistical evaluation of certain parameters or various characteristics. In this case, algorithms for nonparametric estimation of the regression function were used according to observations (3,4).

The computational experiment was carried out as follows:

1. A training sample of regression of the type $y = sin(x)^2$ dependence was formed. Emissions were artificially added when forming the training set.

2. Select the bell-shaped function, set the blur parameter. The triangular core is used as a bell-shaped function:

$$\Phi(v) = \begin{cases} 1 - |v|, & |v| \le 1, \\ 0, & |v| > 1. \end{cases}$$

3. Determined the exam sample for the implementation of the exam or the exam was conducted in a sliding mode.

4. Check each sampling point for the quality of estimation.

Next, we work with the entire sample, constructing a function and its restoration we find the accuracy criterion. As a non-parametric estimate accuracy criterion, we use the quadratic criterion:

$$\sigma^2 = \sum_{i=1}^{s} (y_i - y_s(x_i))^2,$$

where y_i– measured variable value y, $y_s(x_i)$ – its nonparametric evaluation (1).

After checking the accuracy criterion, we draw attention to the points that have a large recovery error and they satisfy the criterion (7). The elements of the training sample that meet the requirement:

$$\rho_i > 2\sigma^2, \tag{7}$$

where $\rho_i = (y_i - y_s(x_i)), i = \overline{1,s}$ are selected and removed from the original sample. In the figures, the following notation is used: 1 is a training sample, 2 is a non-parametric assessment. A triangular core was used as a bell-shaped finite function.

We present the results of a numerical study illustrating the efficiency of the algorithm (7). Consider restoring the regression function from observations, which has several overshoots with a sample size of 100.

Figures 5 and 6 show a sample with two overshoots, in the first case there is no interference, and in the second case 10% is added. Figure 7 shows the operation of the algorithm (7).

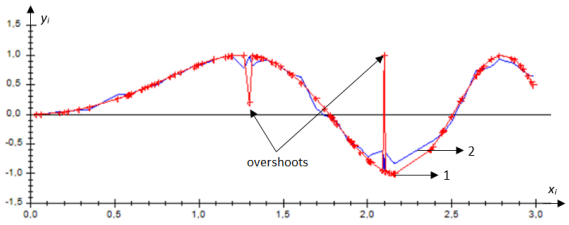


Figure 5: w=0,36

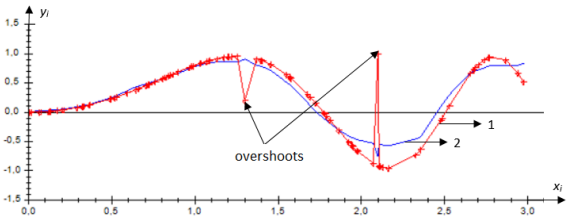


Figure 6: l=10%, w=0,42

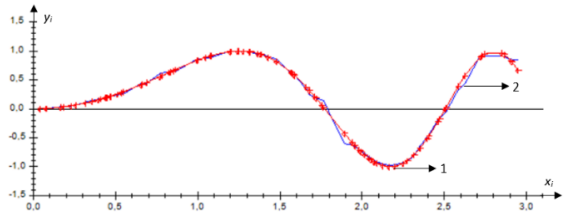


Figure 7: w=0,06

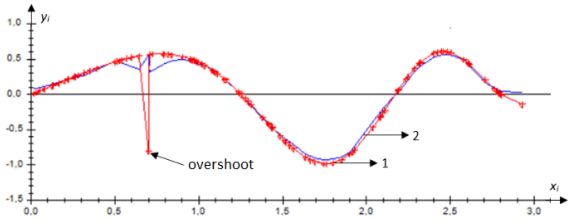


Figure 8: w=0,31

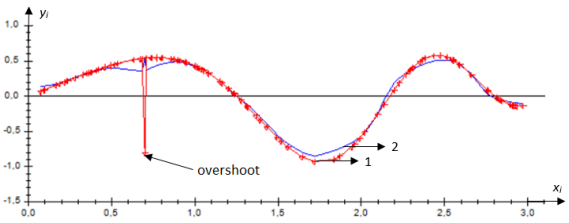


Figure 9: l=10%, w=0,34

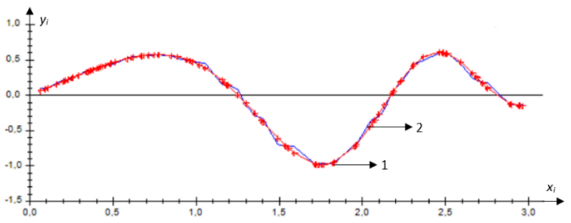


Figure 10: without overshoots, w=0,04

Figures 8 and 9 show the function $y = cos(x)^2 * sin(x)$ with one overshoot. Figure 10 shows the operation of the algorithm (9)

We can conclude about the effectiveness of the proposed algorithms from the analysis of the conducted numerical studies for nonparametric estimation of the regression function from observations with misses.

It is possible to obtain significantly better quality of function recovery from observations using the proposed robust estimation technique. The accuracy of recovery y increased significantly after eliminating overshoots. For clarity, the experiment was considered several functions $y = f(x)$. We also considered cases that correspond to different levels: 0, 5 and 10% on y.

Conclusions

When constructing models of multidimensional inertia-free objects with delay, it is advisable to use the above algorithms for recovering mutually ambiguous characteristics from observations with overshoots. As shown by computational experiments, the accuracy of the predicted variables is significantly increased. Solving such problems required a certain change in the traditional approaches to nonparametric estimation of the regression function, as well as a different approach to estimating $y(x)$ in the presence of overshoots. The restoration of mutually ambiguous characteristics is largely due to the direction of movement along the path chosen by the researcher. it is arbitrary, but it is obligatory to return to the same point from which the movement began.

Nonparametric robust estimation algorithms differ significantly from the generally accepted approach in robust statistics. The main feature of this is the detection and elimination of a slip from the training set.

Both of these tasks are typical in identifying multidimensional inertialess objects. Conducted numerous computational experiments confirm their effectiveness in constructing models of objects. The article presents some typical fragments of the results of extensive numerical studies.

References

[1] Medvedev A.V. (2015). *Fundamentals of the theory of adaptive systems.* SibGAU, Krasnoyarsk.

[2] Eykkhoff P. (1975). *Fundamentals of identification control systems.* Mir, Moscow.

[3] Zhivoglyadov V.P., Medvedev A.V., Tishina E.V. (1973). *Reconstruction of ambiguous static characteristics from experimental data. Automation of industrial experiment.* Frunze.

[4] Korneeva A.A., Chernova S.S., Shishkina A.V. (2017). Nonparametric Algorithms for Recovery Of Mutually Unbeatted Functions on Observations. Ap-

plied Methods of Statistical Analysis. *Nonparametric Methods in Cybernetics and System Analysis – AMSA'2017*. pp. 64-71.

[5] Korneeva A.A., Chernova S.S., Shishkina A.V. (2017). Non-parametric algorithms for recovery of mutually unbeatted functions on observations. *Siberian journal of science and technology*. Vol. **18 №3**, pp. 510-519.

[6] Korneeva A.A., Chernova S.S., Shishkina A.V. (2017). On nonparametric estimation of mutually ambiguous functions from observations. *Automation systems in education, science and production*. pp. 92-98.

[7] Nadaraya E.A. (1983). *Nonparametric estimation of probability density and regression curve*. Publishing house of the Tbilisi University, Tbilisi.

[8] Tarasenko F. P. (1976). *Nonparametric statistics*. Tomsk.

[9] Loner R. L., Uilkinson G. N. (1984). *Sustainable statistical methods for data evaluation*. Moscow.

[10]] Shulenin V. P. (2016). *Robust methods of mathematical statistics*. Tomsk.

[11] Khyuber. P. (1989). *Robustness in statistics*. Mir, Moscow.

[12] Khampel F., Ronchetti E., Raussey P., Shtael V. (1989). *Robustness in statistics. Approach based on influence functions*. Mir, Moscow.

[13] Kitayeva A. V. (2009). *Robust and nonparametric estimation of characteristics of random sequences : thesis*. Tomsk.

[14] Chernova S. S., Shishkina A.V. (2017). On nonparametric estimation of mutually ambiguous functions from observations. *Young scientist. №25*. pp. 13-20.

[15] Sopova, L. N., Chernova S. S. (2017). To the problem of nonparametric robust estimation of the regression function on observations. *Siberian journal of science and technology*. Vol. **18 №4**, pp. 825-832.

Acute pancreatitis severity classification: accuracy, robustness, visualization[1]

Ekaterina Mangalova[2], Olesya Chubarova[3], Daniil Melekh [3] and Anton Stroev[4]
[2] *LLC RD-Science, Krasnoyarsk, Russia*
[3] *Siberian Federal University, Krasnoyarsk, Russia*
[4] *Krasnoyarsk State Medical University named after Prof. V.F.Voino-Yasenetsky, Krasnoyarsk, Russia*
e-mail: e.s.mangalova@hotmail.com

Abstract

The work is devoted to the problem of acute pancreatitis severity classification. This problem is characterized by a small amount of data, which leads to unstable estimations for new patients and a strong influence of the training sample on the predictions. In this paper prediction stability visualization based on violin plot is proposed and applied. A simulation experiments are carried out to study the stability of linear regression, support vector machine, random forest trained with various subsets.

Keywords: classification, machine learning, visualization, violin plot, bootstrapping.

Introduction

Early recognition of disease severity is important to identify patients on admission or during the first 24 to 48 hours who will require aggressive resuscitation. These patients should be treated in an intensive care unit or transferred to a high-acuity care hospital.

Classification of acute pancreatitis defines 3 degrees of severity according to the morbidity: mild, moderately severe, and severe acute pancreatitis.

Mild acute pancreatitis lacks organ failure or local or systemic complications. Pancreatitis resolves rapidly, mortality is rare, pancreatic imaging is often not required.

Moderately severe acute pancreatitis has transient organ failure, local complications, and/or systemic complications but not persistent (>48 hour) organ failure. The morbidity is increased as is mortality (< 8%) compared with that of mild acute pancreatitis.

Severe acute pancreatitis is defined by persistent organ failure and patients usually have 1 or more local and/or systemic complications. Patients with severe acute pancreatitis that develops within the early phase are at a markedly increased risk (36%-50%) of death [1].

[1]The reported study was funded by Krasnoyarsk Regional Fund of Science, to the research project: Development and implementation of decision support system for acute pancreatitis diagnosis and treatment in the Krasnoyarsk Territory

The study was based on a retrospective analysis of 130 cases of acute pancreatitis: 47 cases from Krasnoyarsk Regional Clinical Hospital and 83 cases from RSBHI Regional Interdistrict Clinical Hospital 20 named after I.S. Berzon in the period from 2015 to 2017.

The task is to estimate of acute pancreatitis severity by using patient clinical examination data $D = \{(\bar{x}_i, y_i), i = 1, ..., 130\}$, where $\bar{x} = \{x^1, ..., x^{27}\}$ is set of features (Clinical Blood Analysis, Biochemical Blood Analysis, Ultrasound of pancreas, the results of the examination of the patient) measured in 130 patients.

1 Data preparation

1.1 Feature Scaling

Since the range of values of raw data varies widely, in some machine learning algorithms, objective functions will not work properly without normalization. For example, the Support Vector Machine is based on the distances between points. If one of the features has a broad range of values, the distance will be governed by this particular feature. Therefore, the range of all features should be normalized so that each feature contributes approximately proportionately to the final distance.

All variables are preprocessed using the min-max scaling.

Min-max scaling is the simplest method and consists in rescaling the range of features to scale the range in [0, 1]. The general formula is given as:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

, where x is an original value, x' is the normalized value.

1.2 Filling missing values

Data scientists often check data for missing values and then perform various operations to fix the data or insert new values. The goal of such cleaning operations is to prevent problems caused by missing data that can arise when training a model.

Two types of operations for "cleaning" missing values are implemented:

- Replacing missing values with a linear regression. If two features are strongly correlated linear regression is used to fill missing values. For example, the size of the head, body or tail of the pancreas may be absent due to poor visualization of the pancreas on ultrasound examination of the abdominal cavity. However, the size of the head, body and tail of the pancreas is highly linearly correlated and can be filled.

- Replacing missing values with a within-class median. If features there are not correlated missing values are replaced using a within-class median. This technique allows to avoid reduction of the influence of feature with a large number of missing values as in the case of replacement with median for the whole sample.

2 Accuracy estimation

Since the three classes are strictly ranked, the multi-class classification problem can be solved as a regression problem. As a result, each new object (patient) instead of the class number (1 - mild acute pancreatitis; 2 - moderately severe acute pancreatitis; 3 - severe acute pancreatitis) will be assigned a value from 1 to 3, characterizing not only the class of disease severity, but also how likely this severity class. For example, if the first patient has prediction 1.1 and the second has prediction 1.3, then although they will both be assigned to patients with mild severity of acute pancreatitis, but the probability that the first patient has a mild severity is higher than the second has one.

As accuracy criteria the following indicators were chosen:

- Mean Absolute Error (MAE);
- Mean Squared Error (MSE);
- Correlation Coefficient (Corrcoef);
- Number of Mistakes (NoM). If the prediction differs from the actual value by more than 0.5, it means that the classifier predict wrong class. Such forecasts will be called mistakes.
- Number of Mistakes x2 (NoM x2). If the prediction differs from the actual value by more than 1.5, it means that the classification error is more than one class (mild acute pancreatitis instead of severe acute pancreatitis or vice versa). Such forecasts will be called mistakes x2.

Table 1 contains accuracy of different algorithms calculated using leave-one-out cross-validation technique. Experiments show that SVM provides the greatest accuracy in all indicators.

Table 1: Accuracy of Linear Regression, SVM and Random Forest

	MAE	MSE	Corrcoef	NoM	NoM x2
Linear Regression	0.375	0.269	0.783	44	1
Support Vector Machine	0.354	0.243	0.808	35	0
Random Forest	0.413	0.293	0.765	43	1

3 Robustness

3.1 Small dataset problem

Acute pancreatitis severity classification task is characterized by small sample size for objective reasons. Analysts in medicine face with small dataset problem due to

the prohibition on disclosure and dissemination of personal data. In such tasks, the analyst deals with the following challenges:

- Overfitting. With only a few data, the risk to overfit model is higher;
- Outliers. If analysts have millions of data, a couple of outliers will not be a problem. But with only a few, they will definitely skew prediction results.

The bootstrap procedure [2] can be used to evaluate the robustness of the predictions for the original sample and the effect of certain observations from the initial sample on the predictions.

3.2 Bootstrapping

The basic idea of bootstrapping is that inference about a population from sample data (training set) can be modelled by resampling the sample data and performing inference about a sample from resampled data. As the population is unknown, the true error in a sample statistic against its population value is unknown. In bootstrap-resamples, the 'population' is in fact the sample, and this is known; hence the quality of inference of the 'true' sample from resampled data is measurable.

The bootstrap creates a large number of datasets that we might have seen and computes the statistic on each of these datasets. Thus we get a distribution of the statistic.

In our task, we are interested in the acute pancreatitis severity class of people worldwide. But we cannot measure all the people in the global population, so instead we sample only a tiny part of it, and measure that. Assume the sample (the training dataset) is of size N; that is, we measure the features (Clinical Blood Analysis, Biochemical Blood Analysis, Ultrasound of pancreas, the results of the examination of the patient) of N individuals. From that single sample, only one acute pancreatitis severity prediction can be obtained for each new patient. In order to reason about the population, we need some sense of the variability of the prediction that we have computed.

The most popular bootstrap method involves taking the original data set of N patients and randomly sampling from it to form a new sample (bootstrap sample) that is also of size N. The bootstrap sample is taken from the original by using sampling with replacement. On the first step, we randomly choose N_1 patients from the original data, On the second step, we randomly choose $N - N1$ patients from chosen on the first step. The key parameter for bootstrapping is the ratio between the number of unique observations in the bootstrap sample (N_1) and the initial sample size (N): $p = \frac{N_1}{N}$. This process is repeated a large number of times, and for each of these bootstrap samples we fit model (Linear Regression, Support Vector Machine and Random Forest) and make predictions for new patients.

After applying the bootstrap technique we can have a set of predictions for each new patient that can be analyzed and visualized to make the final decision.

4 Visualisation

4.1 Violin plot

Many different graphs and statistics interpret the characteristics of dataset.

While a box plot [3] only shows summary statistics such as median and interquartile ranges and gives information about location, scale, symmetry and tail thickness, the kernel density estimation shows the full distribution of the data. The difference between the box plot and kernel density estimation is particularly useful when the data distribution is multimodal. In this case a density trace shows the presence of different peaks, their position and relative amplitude.

Violin plots [4] combines the box plot and density trace smoothed by a kernel density estimator and can be used to show robustness of machine learning algorithms.

4.2 Comparison of Machine Learning algorithms

Figure 1 illustrates the influence of the training set on the prediction stability for typical observations from different classes (classes were determined by a medical expert): a - mild acute pancreatitis; b - moderately severe acute pancreatitis; c - severe acute pancreatitis. The ratio p between the number of unique observations in the bootstrap sample (N_1) and the initial sample size (N) is equal to 0.9. The density trace is plotted symmetrically to the upper and the lower of the horizontal box plot. Symmetric plot makes it easier to see the magnitude of the density. The black vertical line shows the median of the predictions, while the gray rectangle depicts interquartile range.

The graph demonstrates ambiguity of severity predictions produced different machine learning algorithms. Note that the Random Forest makes different predictions even with the same training set because of the elements of randomness in the model. When different bootstrap samples are used to fit model, the range of possible forecasts becomes even higher for almost all patients. On the contrary, SVM predicts based on several basic observations. In the case when both bootstrap subsets contain the same basic observations (support vectors), the models trained on them give very close the acute pancreatitis severity estimations. The diversity of SVM forecasts is achieved by subsets that do not contain one or more support vectors.

The Figure 2 shows a comparison of predictions made by different algorithms for patients of the same class (severe acute pancreatitis):

- The predictions of algorithms can be inconsistent, as in the case of Figure 2.a. While Random Forest tends to determine the moderately severe acute pancreatitis, Linear Regression and Support Vector Machine predict a severe acute pancreatitis;

- The predictions of algorithms can be consistent, as in the case of Figure 2.b. This is observed for typical class members for whom the initial training set contains many similar patients.

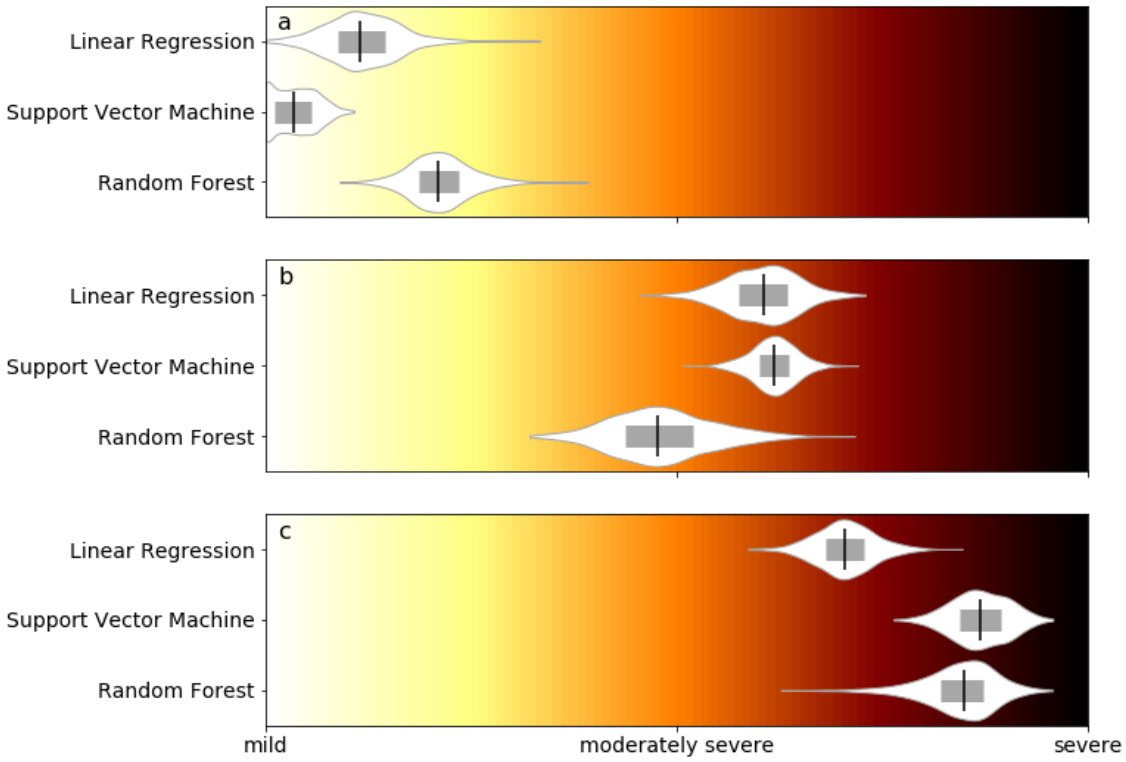


Figure 1: Violin plots based on various model predictions for typical observations from different classes: a - mild acute pancreatitis; b - moderately severe acute pancreatitis; c - severe acute pancreatitis

- The predictions of the algorithms can be incorrect, as in the case of Figure 2.c. Note the large scatter of the random forest predictions to the side of severe acute pancreatitis class that can be interpreted as classifier hesitation.

4.3 The effect of the bootstrap parameter p to the prediction diversity

The ratio between the number of unique observations in the bootstrap sample and the initial sample size p has an impact on predictions. The smaller the value of the parameter p, the smaller the subsets intersect and the greater the differences in the forecasts.

Figure 3 shows the effect of the parameter p on the prediction diversity by the example of one patient. If the parameter p is 0.95, the subsets differ by a maximum of 7 observations and the predictions of the class are compact on the numerical axis. If the parameter p is 0.9, the subsets differ by a maximum of 14 observations, medians change slightly, but the prediction diversity increases significantly for all models. And further, with a decrease in the parameter, this trend continues. When the parameter

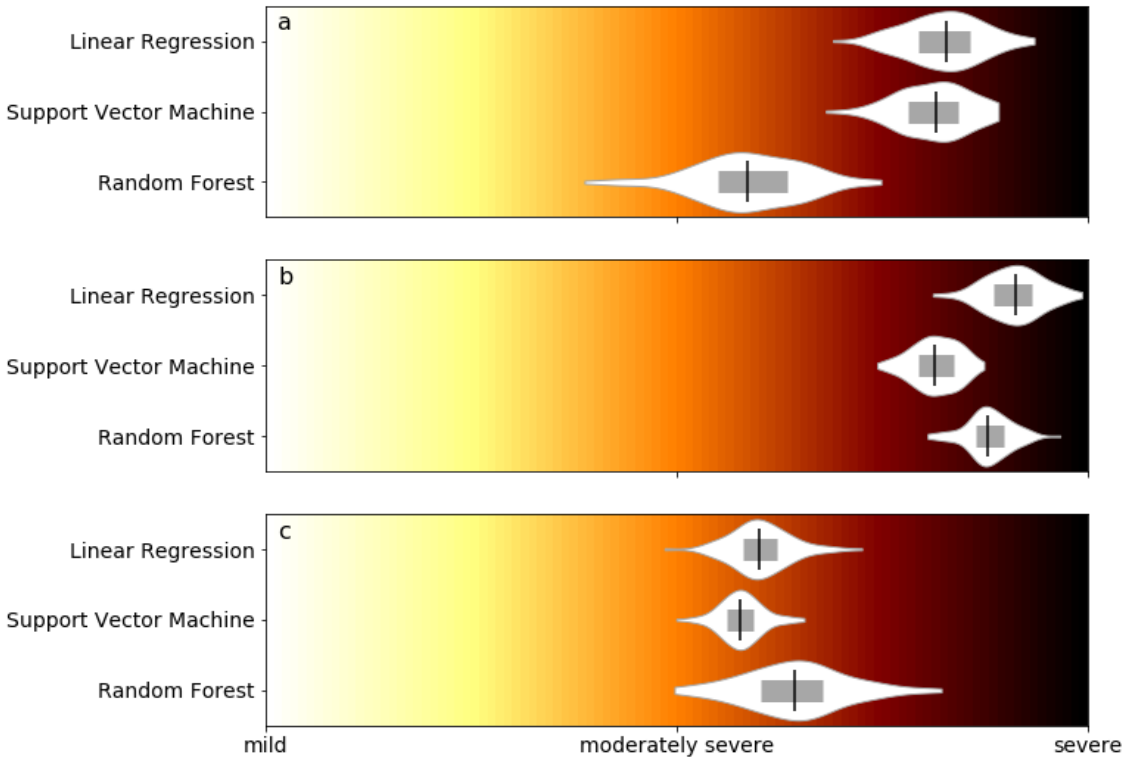


Figure 2: Violin plots based on various model predictions for patients with severe acute pancreatitis

p reaches 0.7, the linear regression and random forest predictions cover almost half of the numeric axis in the range $[1, 3]$.

Taking the final decision on the severity of acute pancreatitis, it is important to consider not only the average value of the forecasts, but also the variance of the forecasts.

Conclusions

Prediction stability visualization procedure was proposed and applied to estimation of acute pancreatitis severity. Visualization method allows to evaluate the prediction diversity of different machine learning algorithms for observation on a single graph. The study compared the stability of forecasts of Linear Regression, Support Vector Machine, Random Forest. This research can be useful to estimate the current dataset quality and to justify the need initial dataset increasing.

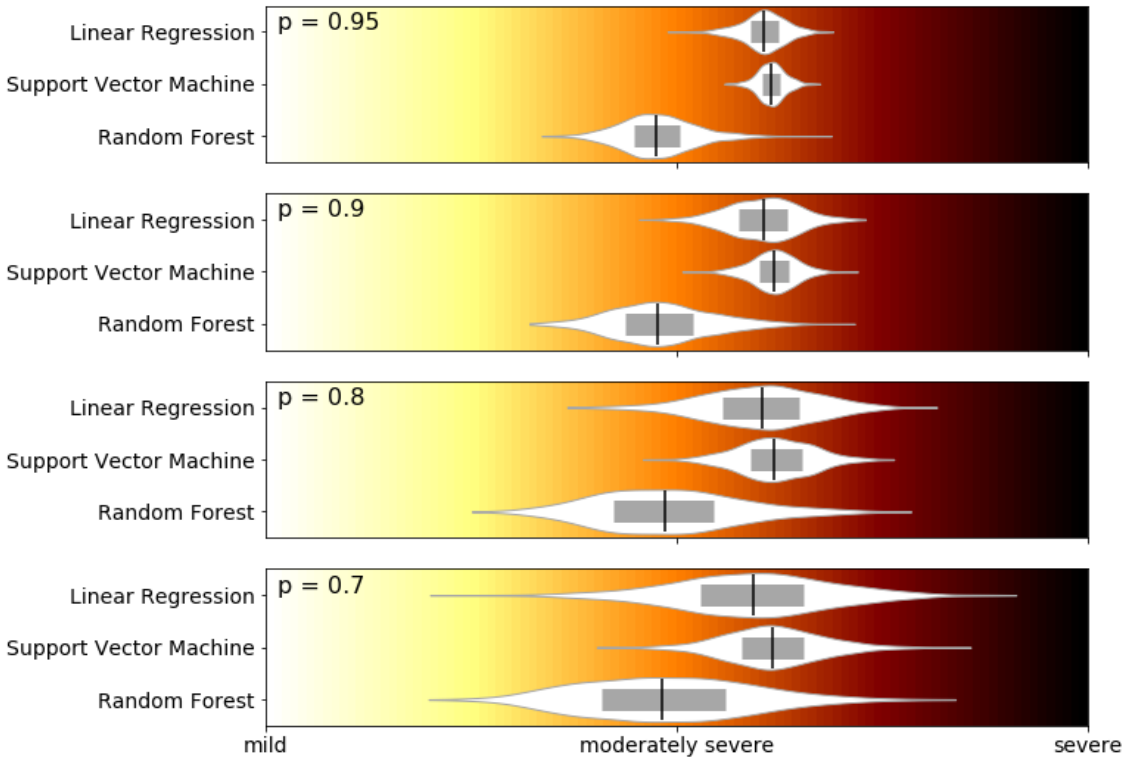


Figure 3: Violin plots based on various model predictions and influence of the ratio p between the number of unique observations in the bootstrap sample and the initial sample size on the stability of predictions

References

[1] Banks P.A., Bollen T.L., Dervenis C., et al. (2013). Classification of acute pancreatitis—2012: revision of the Atlanta classification and definitions by international consensus. *Gut*. Vol. **62**, pp. 102-111.

[2] Efron B., Tibshirani R. J. (1994). An introduction to the bootstrap. – CRC press.

[3] Williamson D.F., Parker R.A., Kendrick J.S. (1989). The box plot: a simple visual method to interpret data. *Annals of internal medicine*. Vol. **110**, pp. 916-921.

[4] Hintze J. L., Nelson R. D. (1998). Violin plots: a box plot-density trace synergism. *The American Statistician*. Vol. **52**, pp. 181-184.

Non-parametric control algorithms for multidimensional H-processes

E.D. Mihov, M.E. Kornet
Siberian Federal University, 79, Svobodny Av.,Krasnoyrsk, Russian Federation
e-mail: edmihov@mail.ru

Abstract

The report is devoted to the control algorithms of inertialess processes. A feature of the problem under consideration is that the components of the input variables vector in the processes under study are in a stochastic dependence. In this regard, the proposed control algorithm, taking into account the specified feature.

Keywords: control algorithms, dual control, non-parametric dual control, H-processes.

Description of the investigated processes

When controlling multidimensional processes, a situation may often occur when the components of the input variables vector are stochastically dependent, and the nature of the relationship between these components is unknown. Such a dependence leads to the fact that the process does not take place in the entire region defined by the input-output variables, but in a certain subdomain. In the following, for reasons of brevity, processes with stochastically dependent input variables will be called H-processes 1. For clarity, we give an example of an H-process with two input variables and one output variable(Figure 1).

The following notation is used in the figure 1: $\Omega^H(\vec{u})$ is the domain of definition of input actions without taking into account the dependence between the input variables; $\Omega^H(\vec{u})$ is the domain of input actions with regard to the relationship between the input variables.

Process under investigation

The article deals with the task of managing multidimensional inertialess H-processes under nonparametric uncertainty. The figure 2 presents the classical object control scheme.

$\vec{u}(t)$ - vector of input controlled actions, of dimension n; $\vec{x}(t)$ is the vector of output variables of the process, of dimension k; $\vec{x}^*$ - vector of defining actions (task); $\xi(t)$ - interference effect on the process.

It is important to note the features that arise when controlling the H-process. As noted above, the process does not take place in the whole area defined by the vector of input and output variables. This means that when choosing a driver action, it is necessary to make sure that the driver action is included in the subdomain of the process. In other words, the driving force is achievable.

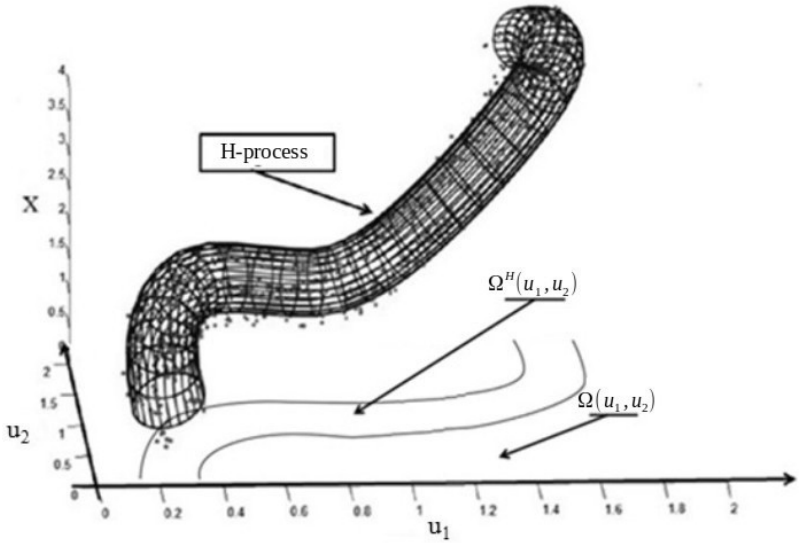


Figure 1: Example of an H-process

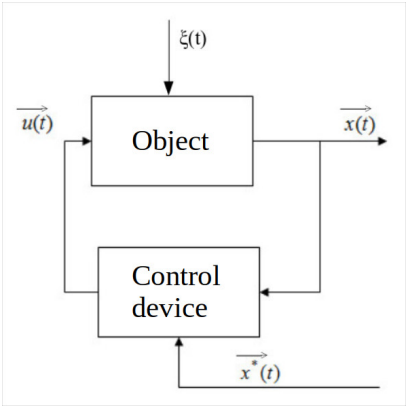


Figure 2: The classical object control scheme

Algorithm of non-parametric dual control

The nonparametric dual control algorithm is an improvement to the algorithm proposed by A.A. Feldbaum [2], which allows researcher to move away from the definition of a parametric model.

It was designed and researched by A.V. Medvedev[1].

The analytical expression of this control algorithm is as follows:

$$\vec{u}_{s+1} = \vec{u}_s(\vec{x}^*) + \delta\vec{u}_{s+1}(\vec{x}^*, \vec{x_s}). \tag{1}$$

$\vec{u}_s(\vec{x}^*)$ – is the component responsible for the accumulation of information about the controlled process (study of the control object). In the initial stages of regulation, this variable is close to zero and does not affect the management of the process,

but with the accumulation of the training sample, its role increases and becomes dominant; $\vec{x}^*$ – setting effect; $\vec{x_s}$ is the sample of observations consisting of the values of the components of the vector of output variables $\vec{x}$.

The component $\vec{u}_s(\vec{x}^*)$ is the mathematical expectation of $\vec{u}$ for a given value of $\vec{x}^*$:

$$\vec{u}_s(\vec{x}^*) = M(\vec{u}(\vec{x}^*)) \tag{2}$$

In the non-parametric dual control algorithm, the non-parametric estimate of the regression function from observations is used as the estimate of $M(\vec{u}(\vec{x}^*))$. [3]:

$$M(u_l(\vec{x}^*)) = \frac{\sum_{i=1}^{s} u_i \prod_{j=1}^{k} \Phi(\frac{x_j^* - x_{i,j}}{c_{sj}})}{\sum_{i=1}^{s} \prod_{j=1}^{k} \Phi(\frac{x_j^* - x_{i,j}}{c_{sj}})}, l = (\overline{1,n}), \tag{3}$$

$\Phi(*)$ – bell-shaped function, $c_{sj}, j = \overline{1,k}$ – blur options.

$\delta\vec{u}_{s+1}(\vec{x}^*, \vec{x}_s)$ – this is the "search step" algorithm. In the initial stages of regulation, this variable makes the main contribution to the management of the process, but with the accumulation of the training sample, its role in management becomes insignificant.

The component $\delta\vec{u}_{s+1}(\vec{x}^*, \vec{x}_s)$ is calculated by the formula 4:

$$\delta u_{s+1,j}(\vec{x}^*, \vec{x}_s) = m \sum_{i=1}^{k} (x_i^* - x_{si}), j = (\overline{1,n}), \tag{4}$$

$(x_i^* - x_{si})$ is difference between the task and the past output of the object, and m is the parameter responsible for the value of the "step"

Impacts produced by the non-parametric dual algorithm have two functions at once: the study of the object and its management. This is the reason for its duality.

Adaptation of non-parametric dual control algorithm for the case of control of the H-process

Note that when controlling a multidimensional H-process, the driving forces determining the desired values of the components of the output vector cannot be chosen arbitrarily, as is customary in control theory. This is due to the fact that it is possible to set such a vector of defining influences that $\prod_{i=1}^{k} \Omega_i^H(\vec{x}^*) = \varnothing$,Ii other words, this action is not achievable for all components of the vector $\vec{x}^*$. In this regard, researcher must first select the achievable (consistent) setting effects $\vec{x}^* \in \prod_{i=1}^{k} \Omega_i^H(\vec{x}^*)$, that is, define $x_1^*, x_2^*, ..., x_k^*$.

The following method of solving the problem is proposed:

1. Calculate value $\sum_{i=1}^{s} \prod_{j=1}^{k} \Phi(\frac{x_j^* - x_{ij}}{c_{sj}})$, $\vec{x}^*$ - defining actions, s - sample size of observations;

2. If the calculated value $\sum_{i=1}^{s}\prod_{j=1}^{k}\Phi(\frac{x_j^*-x_{ij}}{c_{sj}})$ is not equal to 0, it means that the defining influence is achievable, otherwise it is mean that the defining influence cannot be achieved.

In the non-parametric dual control, the search step $\delta\vec{u}_{s+1}$ is calculated using the formula (4). In the case of controlling the H-process with several output variables, the described method for calculating the search step $\delta\vec{u}_{s+1}$ does not fit, since in addition to bringing the object to the task, it is necessary to take into account that the input action must belong to the region $\prod_{j=1}^{k}\Omega^H(\vec{u})$.

Based on the described feature, for calculating $\delta\vec{u}_{s+1}$ it is proposed to use an algorithm with the punishment by randomness.

The modified non-parametric dual control algorithm is as follows:

1) $\vec{u}_s(\vec{x}^*)$ is calculated;

2) a random vector $\delta\vec{u}_{s+1}$ is selected;

3) the value $\vec{u}_{s+1}(x^*)$ is calculated;

4) if $\vec{u}_{s+1}\in\Omega^H(\vec{u})$, then we use $\vec{u}_{s+1}$ as a control, otherwise, return to step 2;

5) if $\sum_{i=1}^{k}|x_{i,s+1}-x_i^*|<\sum_{i=1}^{k}|x_{i,s}-x_i^*|$, then as the next value of the search step $\delta\vec{u}_{s+2}$ it is necessary to take $\delta\vec{u}_{s+1}$, otherwise, again, a random vector is selected as $\delta\vec{u}_{s+2}$;

6) go back to step 1.

The length of the vector $\delta\vec{u}$ is $m|x^*-x_s|$, where m is a customizable coefficient.

The proposed algorithm was used to control the H-process, which has the following structure:

$$\begin{cases} x_1 = u_1 + u_2 + 2 \\ x_2 = -2u_1 + 3u_2 + 1 \end{cases} \tag{5}$$

$$u_2 = sin(\frac{u_1}{2}) + \xi \tag{6}$$

Computational experiments were conducted for 3 cases:

1. the algorithm has no training sample;
2. the algorithm has a small training sample;
3. the algorithm has a large training sample.

The first case in which the size of the training sample s=0 is shown at the figure 3:

As can be seen from the figure, at the beginning the algorithm controls the object rather roughly, but then improves its characteristics due to the experience gained.

The second case, where s=20 is shown at the Figure 4:

As can be seen in the figure 4, the quality of control has improved dramatically, even with a small training sample.

The second case, where s=100 is shown at the Figure 5

As expected, a large amount of the training sample led to precise control of the H-process.

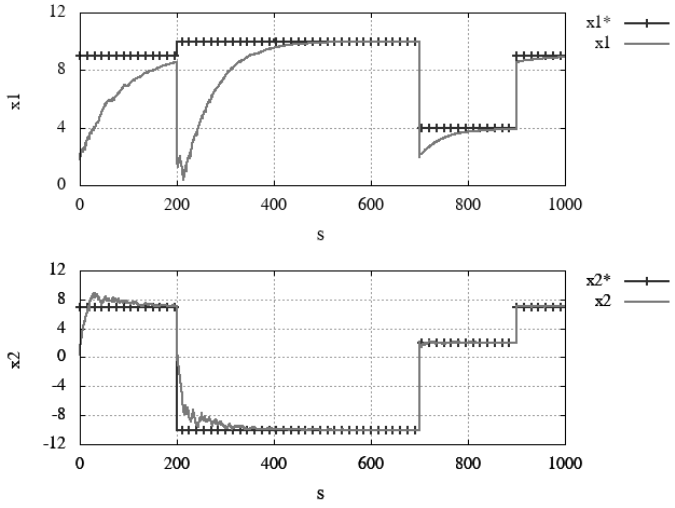


Figure 3: The control process in the case when the algorithm does not have a training sample

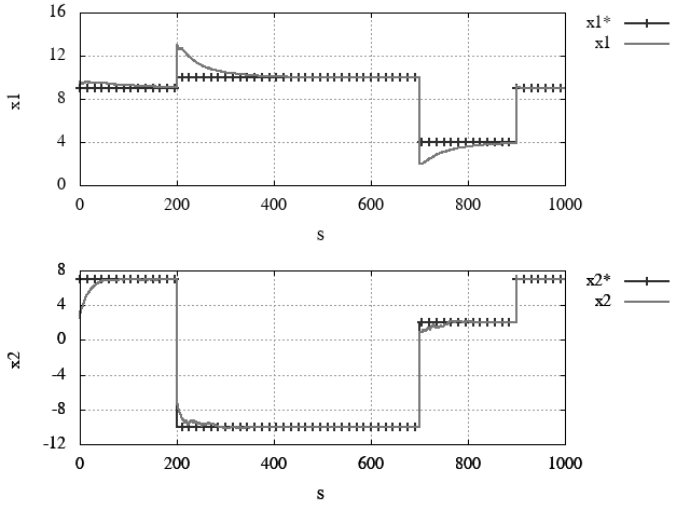


Figure 4: The control process in the case when the algorithm has a small size of training sample

Numerous computational experiments were carried out to control various multidimensional H-processes using the proposed algorithm. Experiments have confirmed that the proposed modification of non-parametric dual control allows researcher to successfully manage the multidimensional H-process.

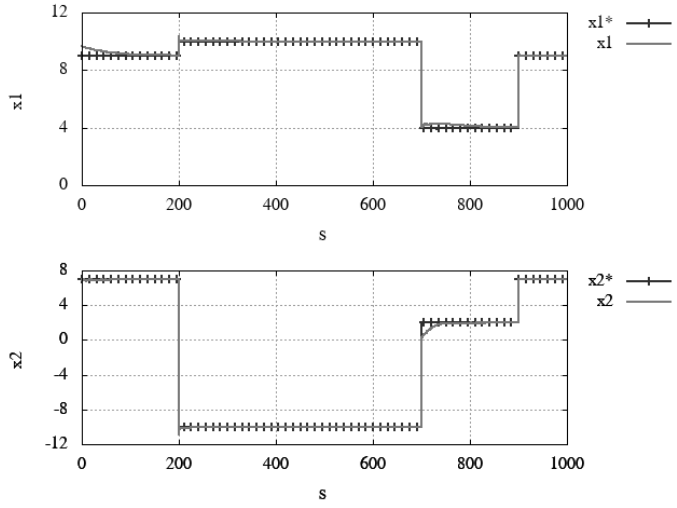


Figure 5: The control process in the case when the algorithm has a large amount of training sample

Acknowledgements

This work was supported by the Ministry of Education and Science of the Russian Federation in the framework of the Federal target program «Research and development on priority directions of development of the scientific-technological complex of Russia for 2014-2020» (agreement № 14.575.21.0142, unique ID project RFMEFI57517X0142)

References

[1] Medvedev A.V. Fundamentals of the theory of non-parametric systems. - Krasnoyarsk: publisher - SibGAU, 2018. - 727 p.

[2] Feldbaum A. A. Fundamentals of the theory of optimal automatic systems. - M.: Fizmatgiz, 1963. - 553 p.

[3] Tsypkin Ya. Z. Adaptation and training in automatic systems. - M.: Science, 1968. - 399 p.

Adaptive algorithm of classification on the missing data

Alexander V. Medvedev[1], Daniil A. Melekh[1],
Natalia A. Sergeeva[2], Olesya V. Chubarova[1]
[1] *Siberian Federal university, Krasnoyarsk, Russia*
[2] *LLC Rd-science, Krasnoyarsk, Russia*
e-mail: danilmelekh@gmail.ru,
n.sergeeva@rd-science.com,
kuznetcova_0@mail.ru

Abstract

The problem of classification by data with gaps, bypassing the stage of their filling, is considered. An adaptive restructuring of algorithms is proposed as a result of the introduction of corresponding indicators into them. The indicators take into account the flow of current information, on the basis of which a decision is made to change the algorithm and the data processing technology itself at each cycle. Computational procedures are based on non-parametric estimation, are given their settings and the results of numerical modeling.

Keywords: supervised learning, missing data, adaptive algorithm, non-parametric estimation of probability density, smoothing window, kernel function, numeric and nominal features.

Introduction

When solving practical problems, the fact of missing values in real data has traditionally been the case. It is possible to solve the problem of processing gaps in the data using different techniques, among which are both suggestions to form training samples only from completely filled objects and fill in the missing values with various approaches and methods. Any initial information about the object is of great value for the researcher, therefore, they most often resort to recovering the missing data, which is already a traditional stage of data preprocessing [1, 3].

The authors of the article were engaged in solving an applied problem related to the classification of objects with a teacher, having a very small sample of data, obtaining of which is also slow. Differences between objects affect the shift of their statistical characteristics in each class. Due to the bias of statistical evaluations of objects of different classes, the choice of tactics for restoring gaps in a training sample for building a classifier and in a new object entering to determine its class is difficult [2, 3].

Also, it was not possible to search for dependencies between features to fill in the gaps due to the smallness of the samples. Therefore, an adaptive classification algorithm was developed that will be able to process data with gaps without a procedure for filling them. The article presents the essence of the algorithm, identifies

the conditions for its use, provides the results of numerical experiments on simulation data and widely known data (Fisher's Iris data set).

The adaptive nature of the classification algorithms is expressed in the reshaping of the training sample from the original, depending on the set of filled features. The computational algorithm changes at each step of the iterative procedure, depending on the completeness of the current information.

To construct adaptive classification algorithms, a modification multidimensional non-parametric probability estimate of the Rosenblatt-Parzen is applied [4].

1 Problem formulation

Let there be a set of objects $\{O_i, i = \overline{1,s}\}$, where s, which are described by a known set of features is the sample size $\{p_i, i = \overline{1,n}\}$, measured in numeric (n_1) and nominal (n_2) scales: $n_1 + n_2 = n$. . For each object there is an indication of the label (i.e., class): $O_i \in Z_l l = \overline{1,L}\}$, L - number of classes. We denote feature measurements for each object with a set of values $\left\{\left(z_i, x_i^j\right), i = \overline{1,s}, j = \overline{1,n}\right\}$, where x_i^j - value of p_j feature at O_i object, z_i - class, s_l - number of class objects Z_l, $\sum_{l=1}^{L} s_l = s$.

Object feature measurements contain omissions. It is necessary to build a classification algorithm that operates with data that contains gaps without a process of their filling, and to develop procedures for setting parameters.

2 Classification algorithm for objects with missing data

For each classified object O_t it is necessary to evaluate its belonging to each class. This procedure involves a comparison with the objects of the training set of each class. The basic idea of the algorithm is to use for evaluation only a set of features O_t that have initial values p. At each t step of the algorithm, the entire training set must be re-sorted relative to the existing set of features p of O_t, presented for classification. Then the size of the training data set may change: $\sum_{l=1}^{L} s_l^t = s^t \leq s$, where s_l^t - class data set Z_l, s^t - total data set after selection of non-empty attributes. Since the set of attributes for assessing the similarity with objects of each class will change, then we denote the number of numeric and nominal features n_1^t and n_2^t, respectively.

Let us demonstrate the idea visually using the example of 3 classes. In the tables below, the filled attributes are highlighted in gray, the features with missing values are displayed without highlighting.

Table 1: Features p of O_t without specifying a teacher

1	2	3	4	5	6	7	8	...	$p_{n_1+n_2}$

Table 2: Initial selection of objects with the teacher

№	Class	1	2	3	4	5	6	7	...	$p_{n_1+n_2}$
1	z_1									
2	z_2									
...	...									
s	z_s									

Subsequent lines of the training set are formed similarly.
For each class, we enter the following value:

$$\alpha_l^t = \sum_{i=1}^{s_l^t} \prod_{i_1=1}^{n_1^t} \Phi\left(\frac{x_t^{j_1} - x_i^{j_1}}{C^{j_1}}\right) \prod_{j_2=1}^{n_2^t} 1(x_t^{j_2}, x_i^{j_2}), \quad l = \overline{1, L}, \tag{1}$$

which constructively repeats the Rosenblatt-Parzen multidimensional nonparametric estimator of probability density. As a bell-shaped function in (1), a triangular kernel, a truncated parabola, a Gaussian kernel, cos, the Sobolev function, and others can be used for features on an numeric scale. For nominal features, the Kronecker delta indicator is used:

$$1(x_t^{j_2}, x_i^{j_2}) = \begin{cases} 1, & \text{if } x_t^{j_2} = x_i^{j_2} \\ \delta_{j_2}, & \text{if } x_t^{j_2} \neq x_i^{j_2} \end{cases}. \tag{2}$$

where δ_{j_2}, $0 < \delta_{j_2} < 1$ — some threshold value for each feature. The specific value of the threshold is selected in the process of training the classifier. Kernel functions satisfy the convergence conditions for nonparametric estimates and are discussed in detail [4].

If $n_1^t = 0$ or $n_2^t = 0$, then the resulting value of the product is limited to some threshold value from the bottom of the entire product, in order to preserve the non-zero value of the other indicators. The more objects differ in the values of the nominal features, the closer the whole work tends to zero, reducing the total weight of the influence of points on the assignment of class values. But at the same time, the difference in only one attribute does not reduce the magnitude of the assessment of the general belonging of an object to a class. A lower bound on the result of the entire work is introduced:

$$0 < \delta < \prod_{j_2=1}^{n_2^t} 1(x_t^{j_2}, x_i^{j_2}) < 1. \tag{3}$$

The smoothing parameter is a vector (by the number of quantitative features). The optimization procedure for setting the parameters of the algorithm is carried out according to the number of points k under the bell-shaped function, the threshold values δ, δ_{j_2}.

The value α_l^t estimates the belonging of an object O_t to a class Z_l according to the initial set. For each new object O_t presented for classification, the volume of the training set for calculating relation (1) will vary due to a different set of unfilled features. Thus, the number of items to be calculated within one class of objects will change. This fact reflects the adaptive nature of the computational procedure, which uses for each object arriving to the classification the newly formed training set from the original.

The next step for deciding whether an object belongs to a particular class is the calculation of normalized estimates β_l^t based on the calculated ones α_l^t:

$$\beta_l^t = \alpha_l^t \Big/ \sum_{l=1}^{L} \alpha_l^t. \tag{4}$$

The probability that an object belongs to a class is proportional to the relative assessment of belonging. The closer it is to 1, the higher the probability of the truth of this class. This can be formulated as follows:

$$O_t \in Z_k | \quad \beta_k^t = \max_{l=\overline{1,L}} (\beta_l^t) \tag{5}$$

As a quality criterion of the classifier, an estimate of the area under the ROC-curve (AUC) is used.

3 Numerical experiment

The numerical study of the algorithm was carried out on three data sets. The first data set corresponded to two non-intersecting classes in the three-dimensional attribute space. The second data set had two intersecting classes in the space of 7 features: 2 in nominal and 5 in numeric scales. To assess the accuracy of the classification, cross-validation using the Monte-Carlo method was used to divide the training and test samples. The calculation results contain the AUC mean value for each sample.

Figure 1: Example of a sample with random gaps

The first sample demonstrates how the algorithm works under favorable conditions. This sample was used with and without gaps. In the case of gaps, two approaches were used: random imitation of gaps (Fig. 1, left side) and random imitation of sequential pairs with disjointed gaps (Fig. 2, right side). The difference of

these approaches is shown in Figure 1. Gaps were introduced in each class in equal numbers, similarly by features.

The share of gaps in both variants was 10% for each attribute, class. Table 1 contains the results of setting the number of points under the kernel function, δ due to the small number of features was set as following: $\delta = \delta_{j_2}^{n_2^t}$, $\delta_{j_2} = 0,01$.

Table 3: First data set, non-intersecting classes

k	AUC			
	No gaps	Random gaps	sequential pairs of disjointed gaps	filled with a median
5	1	0.89	0.9096	1
10	1	0.991	0.994	1
20	1	0.99997	0.9999	1
30	1	0.99956	0.9995	1
Max	1	0.99845	0.9987	1

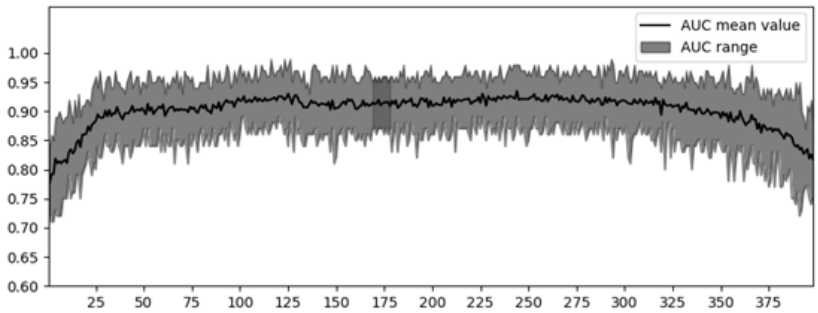


Figure 2: Dynamics of changes in AUC depending on k

The second data set with the intersection of classes contained about 7% of volume of the initial sample of common points in the intersection area. The results of the classification for this sample are presented in Fig. 2, which shows the dynamics of AUC (range and average value) depending on the number of points k under the nuclear function. The average value of accuracy has a stable position, and taking into account the spread of AUC, the best values of k are in the range [169; 178].

The Fischer's Iris appears as the third data set. Table 2 shows the results of the algorithm on accuracy of two Iris classes. Classes with maximum intersection of versicolor and virginica are selected.

To compare the accuracy of the algorithm under the same conditions, in addition to Table 5, the algorithm was tested on a sample with gaps filled with a median by

Table 4: The Fischer's Iris without gaps

k	AUC minimum value	AUC maximum value	AUC mean value
1	0.55	1	0.8358
2	0.59	1	0.84305
3	0.7	1	0.91355
4	0.7	1	0.9141
5	0.7	1	0.9225
6	0.75	1	0.92725
7	0.75	1	0.9338
8	0.8	1	0.9428
9	0.8	1	0.9528
10	0.8	1	0.9563
11	0.8	1	0.95525
12	0.8	1	0.95435

class (Table 6). The filling was done before the stage of dividing the data set into training and test sets, simulating the process of analyzing the labeled data. But in real conditions for newly received unlabeled data it is not possible to use the principle of filling with a median by class.

During the launches of the algorithm, with a gap proportion of 0.1, its accuracy deteriorated by 0.02035 relative to the sample with gaps filled with the median by class and the original sample (Table 6). However, medians for the Fischer's Iris artificially overestimate the accuracy. They contribute to the separation of classes in the original space, because the medians differ from each other. This effect is more clearly observed with an increase in the median filling percentage. Also Table 3 shows the change in the accuracy of the algorithm with different percentage of gaps for each feature. As a result, the algorithm was able to maintain an accuracy of 0.9 under the conditions of 120 gaps from 400 values.

Conclusions

The proposed algorithm can be applied to solve the classification problem if the initial data have gaps. In this case, the algorithm does not require the step of filling the gaps, therefore, its result is not affected by the bias of estimates of the restored gap values. A comprehensive study of the algorithm features requires a comparison with traditional approaches. Also, the further development of the algorithm is supposed to be directed to the formulation of criteria for identifying significant features or their sets for the classification with the teacher in case of having gaps in initial data.

Table 5: First data set, non-intersecting classes

k	Average AUC		
	Percentage of gaps 0.1	Percentage of gaps 0.2	Percentage of gaps 0.3
1	0.70205	0.6639	0.6726
2	0.7374	0.691	0.711
3	0.8847	0.80545	0.80595
4	0.8897	0.81185	0.8336
5	0.9075	0.8474	0.88205
6	0.9136	0.85805	0.8869
7	0.93275	0.87895	0.8969
8	0.9344	0.891	0.89745
9	0.93575	0.9078	0.8987
10	0.93595	0.9155	0.9008
11	0.93785	-	-
12	0.93745	-	-

References

[1] Garcia-Laencin P.J., Sanch-Gomes J.L., Figueiras-Vidal A.R. (2009). Pattern classification with missing data/ *Neural Comput & Applic*, DOI 10.1007/s00521-009-0295-6, Springer-Verlag.

[2] Li Y.Y., Parker L.E. (2008). Classification with missing data in a wireless sensor network *IEEE SoutheastCon 3-6 April*, pp. 533-538.

[3] Zhang S., Jin Z., Zhu X. (2011). Missing data imputation by utilizing information within in-complete instances *Journal of Systems and Software*. Vol. **84**, pp. 452-459.

[4] Härdle W, Linton O. (1994). Applied Nonparametric Methods *Cowles foundation for re-search in economics at Yale university*. 44 p.

Adaptive models for discrete-continuous process

Alina V. Tereshina, Maksim A. Denisov and Daria I. Yareschenko
Siberian Federal University, Institute of Space and Information Technologies, Krasnoyarsk, Russian Federation
e-mail: tereali09@mail.ru, max_denisov00@mail.ru

Abstract

The article is devoted to the construction of a new class of models in the context of lack of priori information. It centers around multidimensional inertialess objects, when the components of the output vector are stochastically dependent, but the nature of this dependence is unknown. In the process of building a model it is necessary to accomplish the solution of implicit functions when the inputs and outputs vectors are non-linear. It should be noted that the form of these functions with an accuracy of the parameters vector is unknown. In this regard, it is necessary to use T-models, when the prediction of output variables is carried out using known input.

Keywords: identification, forecasting, mathematical modeling, T-models, T-processes.

Introduction

Identification of multidimensional stochastic processes is a topical issue for many technological processes of discrete-continuous nature. In real practical problems, such as in [1], the discreteness of object's output variables control can be performed differently. For example, some indicators are measured electrically and others by physico-mechanical or laboratory tests.

A discrete-continuous process is a process that is continuous, but the data is recorded in a discrete time interval. This leads to the fact that dynamic processes, by their nature, are forced to be considered as inertialess with delay.

This article centers around the problem of identifying processes whose output variables are stochastically dependent, however, that dependence is unknown. It is called T-processes [2] and the identification problem in this case is to build T-models of multidimensional statistical objects.

It should be mentioned that the term "process" is not considered as a process of probabilistic nature (for instance, stationary, Gaussian, Markov etc. [3]). Below we discuss the T-processes which occur or develop in time, for instance, technological or economical ones. These processes were first mentioned by A. V. Medvedev in [4].

1 T-model

The system of equations describing T-processes in general form [5] can be represented as follows:

$$F_j(u(t), x(t)) = 0, \quad j = \overline{1, n}, \tag{1}$$

where $u(t)$ is an input vector of variables, $x(t)$ is an output vector of variables. Although, in practice, it is more likely to be a situation when the system of equations (1) can be represented in the following form:

$$F_j(u^{<j>}(t), x^{<j>}(t)) = 0, \quad j = \overline{1,n}, \tag{2}$$

where $u^{<j>}(t), x^{<j>}(t)$ are composite vectors. Composite vector is the vector which includes some components, for instance: $x^{<j>}(t) = (u_2(t), u_5(t), x_2(t), x_7(t))$. At the same time, the main feature of modeling such process under nonparametric uncertainty is the fact that the type of functions in (1) is unknown. In this case, the system of equations takes the form:

$$\hat{F}_j(u^{<j>}(t), x^{<j>}(t), \vec{x}_s, \vec{u}_s) = 0, \quad j = \overline{1,n}, \tag{3}$$

where $\vec{x}_s, \vec{u}_s$ are time vectors (data set received at the s-th time point), for example: $\vec{x}_s = (x_1, \ldots, x_s) = (x_{11}, x_{12}, \ldots, x_{1s}, \ldots, x_{21}, x_{22}, \ldots, x_{2s}, \ldots, x_{n1}, x_{n2}, \ldots, x_{ns})$. However, functions $\hat{F}_j(\cdot)$, $j = \overline{1,n}$ remain unknown for this case. In the theory of identification, such problems are not only considered, but even not defined. Most often it is followed the path of choosing the parametric structure (1) but overcoming this stage is difficult due to the lack of a priori information [6]. In addition, it takes a long time to determine the parametric structure, i.e. model representation in the form of:

$$\hat{F}_j(u^{<j>}(t), x^{<j>}(t), \alpha) = 0, \quad j = \overline{1,n}, \tag{4}$$

where α is parameter vector.

Further, there is a procedure for estimating the parameters using the training sample u_i, x_i, $i = \overline{1,s}$ and the subsequent solution of the nonlinear interrelated ratios system. The success of building a model in this case depends on the qualitative parametrization of the system (4) [7].

2 Computational experiment

For the computational experiment we have multidimensional object with five input variables $u(t) = \{u_1(t), u_2(t), u_3(t), u_4(t), u_5(t)\}$, which take random values in the interval $u(t) \in [0;3]$ and with four output variables $x(t) = \{x_1(t), x_2(t), x_3(t), x_4(t)\}$ which take their values in the following intervals: $x_1(t) \in [-0.54; 18.23]$, $x_2 \in [-0.58; 35.9]$, $x_3 \in [-2.27; 60.5]$, $x_4 \in [-3.23; 75.73]$.

It is formed a sample of input and output variables for the object based on the system of equations, which are chosen arbitrarily (under a computational experiment):

$$\begin{cases} x_1(t) - 2u_2(t) - u_5(t) - 0.3x_2(t) = 0; \\ x_2(t) - u_1^3(t) - 0.3u_3(t) - 0.5x_1(t) = 0; \\ x_3(t) - u_4^2(t) - \sqrt{u_5(t)} - 0.2x_4(t) = 0; \\ x_4(t) - u_2^2 - u_3(t) - 0.4x_3(t) = 0. \end{cases} \tag{5}$$

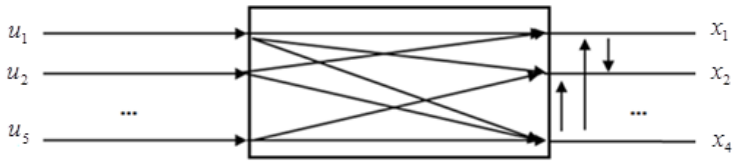


Figure 1: Multidimensional object

The system of equations (5) is not a description of the real process and is introduced only for a given computational experiment. Such representation allows to conduct a modeling process and compare the results of output vector components estimates which are obtained using the T-model. In examine real-world problems the training and test samples are obtained by conducting a series of numerous experiments with the studied object.

At the first stage of the computational experiment we solve the system (5). To do this we define it relative to $x(t)$ with known values of the $u(t)$, which can be formed randomly from the mentioned above intervals.

In the form of qualitative dependence system (5) can be represented as follows:

$$\begin{cases} \hat{F}_{x_1}(x_1(t), x_2(t), u_2(t), u_5(t)) = 0; \\ \hat{F}_{x_2}(x_1(t), x_2(t), u_1(t), u_3(t)) = 0; \\ \hat{F}_{x_3}(x_3(t), x_4(t), u_4(t), u_5(t)) = 0; \\ \hat{F}_{x_4}(x_3(t), x_4(t), u_2(t), u_3(t)) = 0. \end{cases} \tag{6}$$

Thus, it is necessary to estimate the values of the output variables using the known components of the input $u(t) = \{u_k(t), k = \overline{1,5}\}$. This is the main result of the identification problem solution. Surely, one is tempted to name the system of equations (6) as a model of the process under study, but the reality is different due to the functions $F(x)$ are unknown. That is why the chain of corresponding non-parametric statistics acts as the T-model.

At the first stage of the experiment it is calculated the residual errors for each component of the output vector using the following formula:

$$\varepsilon_j(i) = F_{\varepsilon j(u^{<n>}, x_j(i))} = x_j(i) - \frac{\sum\limits_{i=1}^{s} x_j[i] \prod\limits_{k=1}^{<n>} \Phi\left(\frac{u_k^{'} - u_k[i]}{c_{su_k}}\right)}{\sum\limits_{i=1}^{s} \prod\limits_{k=1}^{<n>} \Phi\left(\frac{u_k^{'} - u_k[i]}{c_{su_k}}\right)} \tag{7}$$

where $j = \overline{1,n}$ and $< m > \leq m$ is the number of dimensions of the composite vector u_k.

Bell-shaped functions $\Phi\left(\frac{u_k^{'} - u_k[i]}{c_{su_k}}\right)$ and bandwidth parameter c_{su_k} satisfy convergence conditions [8] and have the following properties:

$$\Phi(\cdot) < \infty;$$

$$c_s^{-1} \int\limits_{\Omega(u)} \Phi(c_s^{-1}(u-u_i))du = 1;$$

$$\lim_{s\to\infty} c_s^{-1}\Phi(c_s^{-1}(u-u_i)) = \delta(u-u_i);$$

$$\lim_{s\to\infty} c_s = 0;$$

$$\lim_{s\to\infty} sc_s = \infty.$$

In this experiment, the triangular core was chosen as a bell-shaped function.

It should be added that the residual errors can be represented in the form which is mentioned below:

$$\begin{cases} \varepsilon_1(i) = \hat{F}_{x_1}(x_1^i(t), x_2^i(t), u_2^l(t), u_5^l(t)); \\ \varepsilon_2(i) = \hat{F}_{x_2}(x_1^i(t), x_2^i(t), u_1^l(t), u_3^l(t)); \\ \varepsilon_3(i) = \hat{F}_{x_3}(x_3^i(t), x_4^i(t), u_4^l(t), u_5^l(t)); \\ \varepsilon_4(i) = \hat{F}_{x_4}(x_3^i(t), x_4^i(t), u_2^l(t), u_3^l(t)). \end{cases} \tag{8}$$

Accordingly, each residual error is compliant to a specific output of the object.

In order to understand how accurate the model corresponds to the object, it is necessary to calculate the error value, which mathematical description presented below:

$$\delta_j = \frac{\sum\limits_{j=1}^{s} |x_j(t) - x_j(u^{<j>}(t))|}{\sum\limits_{j=1}^{s} |x_j(t) - M_{x_j}|} \tag{9}$$

In the computational experiment it is calculated the optimal bandwidth parameter. Sample sizes are $s = 1000$ and $s = 3000$. The results are presented in the figures 2, 3, 4, 5 and the table 1 below.

Table 1: The error values (9) for different noise levels and sample sizes

	$s = 1000$	$s = 3000$
without noise	$\delta_1 = 0.01$	$\delta_1 = 0.002$
5% noise	$\delta_1 = 0.264$	$\delta_1 = 0.261$
10% noise	$\delta_1 = 0.431$	$\delta_1 = 0.337$

Figures 2 and 3 show the true values of the output variables and their prediction. As it can be seen, the process of modeling gives quite accurate result (according to values in table 1 without noise). This indicates the high accuracy of the T-model.

In the second computational experiment we show how the 5% and 10% noise levels, which are imposed on the output component values, affect the final simulation

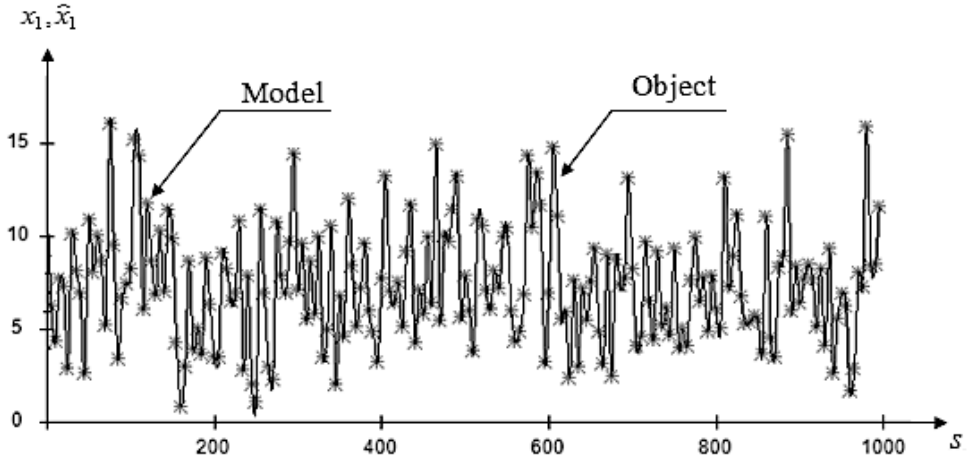


Figure 2: Output variable x_1 model values without noise, $s = 1000$, optimal $c_s = 0.3$

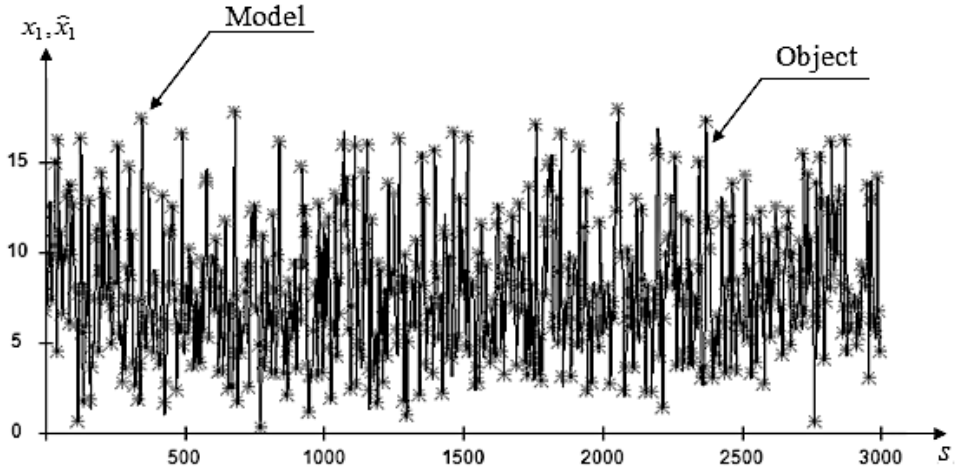


Figure 3: Output variable x_1 model values without noise, $s = 3000$, optimal $c_s = 0.3$

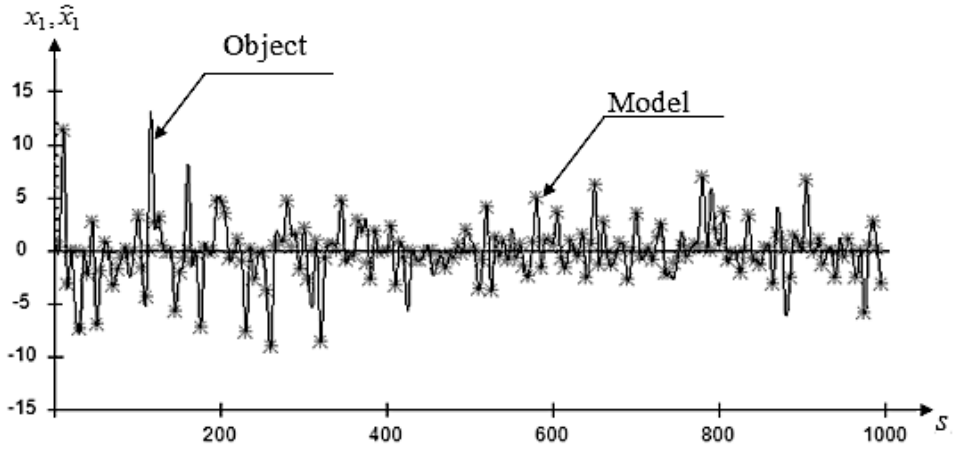


Figure 4: Output variable x_1 model values with noise equals 5%, $s = 1000$, optimal $c_s = 0.3$

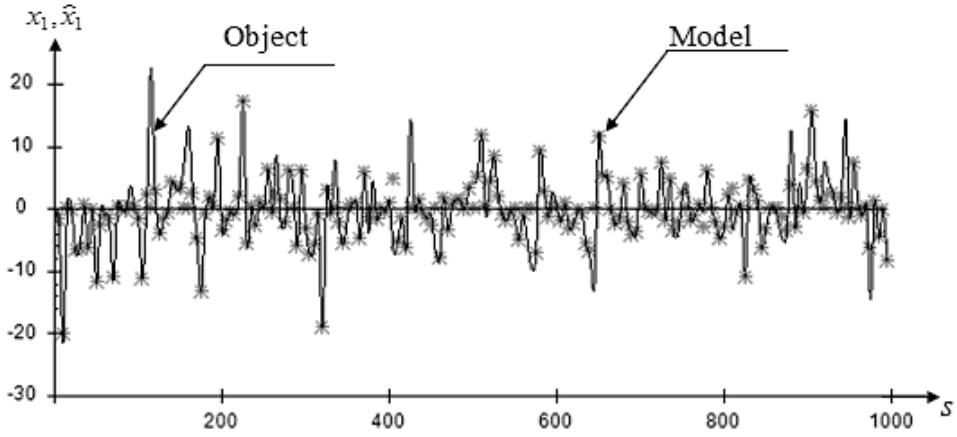


Figure 5: Output variable x_1 model values with noise equals 10%, $s = 1000$, optimal $c_s = 0.3$

results. The sample size is set to $s = 1000$ and $s = 3000$. We also give graphs for one of the outputs as it has been done above (figures 4 and 5).

Figures 4 and 5 show the true and predicted values of the output variables with different noise levels. Results in table 1 and figures 4, 5 show how accurate the model approximates the object.

Conclusions

The article has been considered the problem of identifying inertialess multidimensional objects with delay as well as unknown stochastic connections of the components of the output vector. The experiments described in the paper include a change in a sample size. In addition, the output variables of the model were subjected to stochastic noise of various levels. Computational experiments have shown the high efficiency of the T-modeling process according to values of modeling mistakes.

References

[1] Denisov M.A., Chzhan E.A., Korneeva A.A. (2019). Non-Parametric Approach for Preliminary Processing of Earth Remote Sensing Data. *E3S Web of Conferences* . Vol. **75**, p. 01015.

[2] Medvedev A.V. (1983). *Neparametricheskie sistemy adaptacii [Nonparametric adaptation systems]*. Nauka Publ., Novosibirsk.

[3] Doob J.L. (1953). *Stochastic Processes*. Jhon Wiley & Sons, New York.

[4] Medvedev A.V. (2015). *Osnovy teorii adaptivnyh system [Fundamentals of adaptive systems theory]*. SibGAU Publ., Krasnoyarsk.

[5] Medvedev A.V. (2018). *Osnovy teorii neparametricheskih sistem. Identifikaciya, upravlenie, prinyatie reshenij [Fundamentals of the theory of nonparametric systems. Identification, control, decision making]*. SibGAU Publ., Krasnoyarsk.

[6] Eickhoff P. (1974). *System Identification: Parameter and State Estimation*. Wiley & Sons, London.

[7] Ljung L. (1999). *System Identification: Theory for the User*. Prentice-Hall, New Jersey.

[8] Nadaraya E.A. (1963). On estimating regression. *Theory of Probability and Its Applications* . Vol. **10**, p.p. 186-96.

Nonparametric dual control algorithm for discrete linear dynamic systems

A.V. Raskina[1], E. A. Chzhan[1], V.V. Kukartsev[1,2], A.V. Karavanov[1], A.V. Lonina[1]
[1] *Siberian Federal University, Krasnoyarsk, Russia*
[2] *Siberian State University of Science and Technology, Krasnoyarsk, Russia*
e-mail: raskina.1012@gmail.com

Abstract

The article devotes to the problem of controlling discrete linear dynamic systems under non-parametric uncertainty. Control action are calculated in which the difference equation degree of a dynamic process model is refined based on the rule of selection of significant variables. The computational experiments confirmed the efficiency of using non-parametric algorithms to control dynamic systems in comparison with the PID algorithm and the quasi-optimal control system.

Keywords: non-parametric algorithm, discrete dynamic system, priory information.

Introduction

Designing intelligent systems for controlling dynamic objects is one of the important task of system analysis. Previously, algorithms of control dynamic objects were developed, in particular, the most widely used standard control algorithms. In some cases, their use is not effective enough. Into contemporary scientific approaches, optimal control algorithms are used. However, for their application, as a rule, a priori knowledge of the structure and parameters of the controlled object is necessary.

In the conditions when there is no prior information, the development of new control algorithms is a significant scientific problem. One of the ways to solve this task to use non-parametric methods [5]. For the application of non-parametric methods, it is necessary to know only about the quality characteristics of the object under study.

1 Dual control

Dual control was suggested by A.A. Feldbaum [7] and developed on the basis of the theory of statistical solutions. The theory of dual control was further developed in the studies of various authors [4], in particular B. Wittenmark [2]. It should be noted that a system in which dual control algorithms are used is an adaptive system, because as current information is received from an object, the quality of functioning increases.

2 Problem set-up and algorithm

The paper considers classes of control objects that can be described by linear difference equations of the form (1).

$$x_t = F(x_{t-1}, ..., x_{t-k}, u_t, \xi_t). \tag{1}$$

where F is an unknown linear functional, k is the degree of a difference equation, which is limited $k \leq k_{max}$. The input u_t and output x_t of a dynamic object are represented by measurements that form a sample of the form $x_i, u_i, i = \overline{1,s}$, where s is the sample size, u_i, x_i are the measurements of the input and output of the object at a time instant t_i.

For a dynamic object that can be described by difference equation (1), the control problem is to find the control functions u_t. The control function translates the output of an object x_t to a specified value x_t^* in some finite time t_p. In this case, the functional F is assumed to be unknown from a priori information, but there is a sample of observations $x_i, u_i, i = \overline{1,s}$. Non-parametric dual control algorithm has the following form (2) [3]:

$$u_{s+1} = u_s^* + \Delta u_{s+1}, \tag{2}$$

where u_s^* is the component accumulating information about the object under study, $\Delta u_{s+1} = \varepsilon(x_{s+1}^* - x_s)$ is the "studying" search step. The dual control scheme is shown in Figure 1.

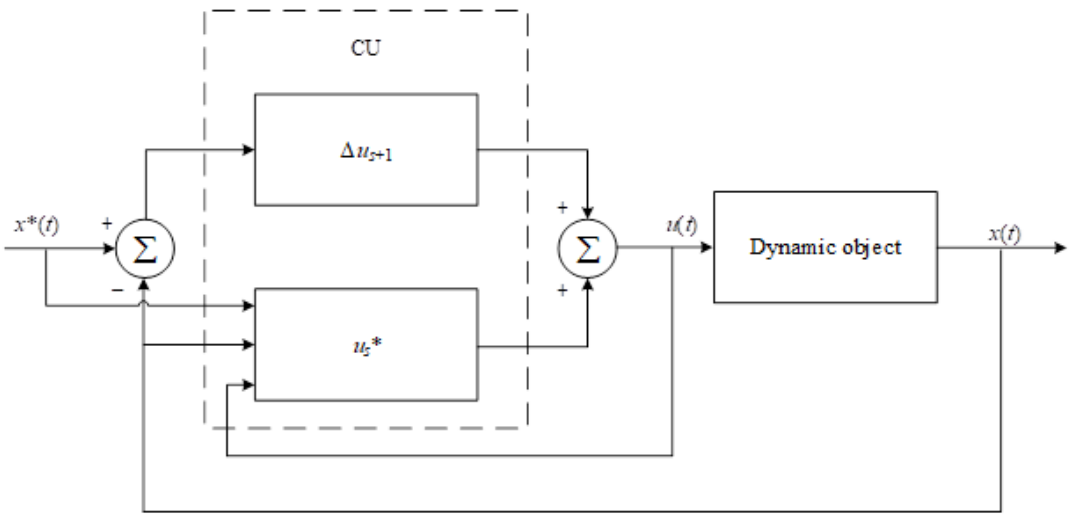


Figure 1: Dual control scheme of a dynamic object

In this case, we use the following estimation to get the value u_s^* from equation (3):

$$u_s^* = \frac{\sum_{i=1}^{s} u_i \cdot \Phi\left(\frac{x_{s+1}^* - x_i}{c_s}\right) \cdot \prod_{j=1}^{k}\left(\frac{x_{s-j} - x_{i-j}}{c_s}\right)}{\sum_{i=1}^{s} \Phi\left(\frac{x_{s+1}^* - x_i}{c_s}\right) \cdot \prod_{j=1}^{k}\left(\frac{x_{s-j} - x_{i-j}}{c_s}\right)}, \quad (3)$$

where u_s^* is a kernel function, c_s^u, $c_s^{x[j]}$ are bandwidths. The optimal bandwidths are found by minimizing a quadratic error function by using the sliding exam method.

3 Essential variables

The control algorithm for dynamic systems is constructed as follows. The differential equation degree of the dynamic process model k is determined on the bases of the rule of selection of essential variables. The vale k is further used in the calculation of control actions in (3), where only selected variables are present.

Formulation of the rule: in formula (3), each variable $x_{s-1}, ..., x_{s-k}$ is assigned its own bandwidths $c_s^{x[1]}, ..., c_s^{x[j]}$, the greater bandwidths, the less influence this variable has on the output of the object.

The algorithm for calculating significant variables x_{i-j} is based on the following scheme. First, the initial value of k is given. The model is constructed by equation.

$$x_s^t = \frac{\sum_{i=1}^{s} x_i \cdot \Phi\left(\frac{u_s - u_i}{c_s^u}\right) \cdot \prod_{j=1}^{k}\left(\frac{x_{s-j} - x_{i-j}}{c_s}\right)}{\sum_{i=1}^{s} \Phi\left(\frac{u_s - u_i}{c_s}\right) \cdot \prod_{j=1}^{k}\left(\frac{x_{s-j} - x_{i-j}}{c_s}\right)}, \quad (4)$$

and the relative error W_0 is calculated:

$$W = \sqrt{\frac{1}{s}\sum_{i=1}^{s}(x_i - x_i^s)^2 \Big/ \sum_{i=1}^{s}\frac{1}{s-1}(m_x - x_i)^2} \quad (5)$$

where m_x is an expected value.

For each i - th iteration, the following set of actions is performed:

1. For each coefficient $c_s^{x^1}, ..., c_s^{x^k}$ the optimal value is found: $c_s^{x^1} = c_s^{*x^1}$, $c_s^{x^2} = c_s^{*x^2}$,...,$c_s^{x^k} = c_s^{*x^k}$;
2. The maximum of all the values obtained is found: $c_{max_s}^{x^j}$;
3. The model is constructed by the equation (4). The multiplier $\Phi\left(\frac{u_s - u_i}{c_s^u}\right)$ is excluded, taking into account that j is a number for $c_{max_s}^{x^j}$.
4. A relative error W_i is calculated.

These actions will be repeated until $W_i \geq W_{i-1}$.

We choose a non-gradient multidimensional optimization the Nelder-Mead method as an optimization algorithm, since this method is effective at a low speed of calculation of the minimized function. To select the initial vertices of a deformable

polyhedron, a region of possible values of bandwidths of the kernel functions was set, from which $n + k + 1$ points were chosen arbitrarily, where n is the number of input variables, k is the degree of the difference equation, which form the simplex $n + k$.

In the case of the relation of an object to a class of linear, the algorithm allows one to determine the structure of the model with an accuracy of parameters.

4 Computational experiment

To illustrate the performance of the proposed algorithm, an example will be considered.The results of controlling a dynamic object using a non-parametric dual control algorithm (3) were compared with the control results of a typical control algorithm (PID), and with the results of using an quasi-optimal control system (criterion is control time).

The control quality was estimated by the control time (t_p) —the time from the beginning of the control to the moment when the output quantity differs from of no more than some given value α. ($\alpha = 5\%$). As an example, we give the work of three control algorithms. The control object is a series connection of three aperiodic links. Detailed control results are shown in table.

Table 1: Comparison results of the non-parametric dual control algorithm (NDCA) with the PID controller and the quasi-optimal control system (QOCS)

	Type of control systems	t_p
1	PID	7.9
2	NDCA	8.1 (at the stage of information accumulation) 1.2 (after passing the stage of accumulation of information)
3	QOCS	3.5

The dualism of the algorithm (3) is as follows. At the first control cycles, the main role in the formation of control actions is played by the term Δu_{s+1} from formula (3). But with the accumulation of information about the object, the role of the term u_s^* increases. Thus, the use of the non-parametric algorithm after passing through the stage of accumulation of information alows it possible to reduce the control time under equal conditions of noise and sample size compared to a typical PID controller and a quasi-optimal control system.

Conclusions

Nonparametric dual control algorithms for dynamic objects have been developed. A distinctive feature of the algorithms is the use of information about the order of a difference equation of a dynamic object when calculating control actions. The task of controlling dynamic objects is most effectively solved by the proposed algorithm,

as compared with typical control algorithms, in particular, the PID algorithm and the quasi-optimal control system.

References

[1] Medvedev A.V. (2015). *Fundamentals of the theory of adaptive systems*. Pub. SibSAU, Krasnoyarsk.

[2] Feldbaum A.A. (1963). *Fundamentals of the theory of optimal automatic systems*. Pub. Fizmatgiz, Moskow.

[3] Lia Duan, Qianb Fucai, Fuc Peilin. (2008). Optimal nominal dual control for discrete-time linear-quadratic Gaussian problems with unknown parameters. *Automatica* Vol. **44, Issue 1**, pp. 119–127.

[4] Wittenmark B. (1995). Adaptive dual control methods: An overview. *In 5th IFAC Symposium on Adaptive Systems in Control and Signal Processing* pp. 67–72.

[5] Medvedev A.V. (2013). Theory of non-parametric systems. Control - I *Vestnik of the Siberian State University named after Academician M.F. Reshetnev* Vol. **2 (48)**, pp. 57-63.

Numerical study of the bioclimatic index of severity of climatic regime based on a stochastic model of the joint meteorological time series

Marina S. Akenteva[1], Nina A. Kargapolova[1,2]
and Vasily A. Ogorodnikov[1,2]
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia*
e-mail: marina-akenteva@mail.ru, nkargapolova@gmail.com, ova@osmf.sscc.ru

Abstract

In this paper, a numerical stochastic model of the joint non-stationary time-series of the wind speed modulus, air temperature and relative humidity is proposed. It is shown that this model may be used for studying the statistical properties of the time series of the bioclimatic index of severity of climatic regime.

Keywords: stochastic simulation, periodically correlated random process, meteorological time series, bioclimatic index of severity of climatic regime.

Introduction

Air temperature and relative humidity, wind speed and atmospheric pressure, as well as other meteorological parameters significantly affect the state of a human being and his or her ability to work. To describe the combined effects of various meteorological parameters, different climatic indicators and weather indices are used. Some of them, such as the heat index, are used to describe the effects of high temperatures and of a high relative humidity [1]. Other indices, such as the Siple index and the Hill wind chill index, are used to characterize the effects of low temperatures and high winds [9, 17]. One of the most universal indicators is the bioclimatic index of severity of climatic regime (BISCR), proposed in [5, 6]. In this paper, probabilistic properties of the time series of the BISCR are studied, as well as the possibility of constructing an appropriate stochastic model.

1 Bioclimatic index of severity of climatic regime

In this paper, the probabilistic properties of the time series of the bioclimatic index of severity of climatic regime (BISCR) are investigated. According to [7], the BISCR is an integral indicator to the degree of bioclimatic discomfort in various types of the vital activity. The BISCR (a non-dimensional value) is a function of a number of meteorological and physiographic parameters, including air temperature (measured in

degrees Celsius), atmospheric pressure (in hectopascals), wind speed modulus (in meters per second), relative humidity of the air (measured in percentage), and elevation of the terrain under consideration above the sea level (in meters).

To calculate the BISCR, the following formula is used [3, 12]:

$$B = \frac{\tilde{T}(P-266)(1-0.02V)}{75\tilde{R}\tilde{S}},$$

where $\tilde{T}, \tilde{R}, \tilde{S}$ are the temperature, humidity and radiation coefficients, P, V are the atmospheric pressure and the wind speed modulus, respectively. The coefficients are determined by the following formulas:

$$\tilde{T} = \begin{cases} 1 - 0.0089\,(22 - T)\,, & \text{if } T < 22, \\ 1 - 0.0263\,(T - 22)\,, & \text{if } T \geqslant 22, \end{cases}$$

$$\tilde{R} = \begin{cases} 1 + 0.06\frac{(50-R)}{100}, & \text{if } R < 50, \\ 1 + 0.06\frac{(R-50)}{100}, & \text{if } R \geqslant 50. \end{cases}$$

Here T, R are air temperature and relative humidity. The radiation coefficient $\tilde{S}$ depends only on the elevation H of the terrain above the sea level:

$$\tilde{S} = \begin{cases} 1, & \text{if } H < 2000, \\ 1 + 0.45\frac{(H-2000)}{1000}, & \text{if } H \geqslant 2000. \end{cases}$$

We should note that as the basis for the creation of the BISCR are empirical concepts of "comfort" and "extremes" of the influence of meteorological factors on a human being, as well as the condition that the index equals 10 under special meteorological and geographical conditions, i.e. at $T = 22$ °C, $R = 50$ %, $V = 0$ m/s, $P = 1016$ hPa, $H < 2000$ m [3, 7].

2 A stochastic model

We have to simulate the time series $\vec{B} = (B_1, ..., B_8, B_9, ..., B_{8N})$ of the BISCR on N days interval with 8 measurements per day.

There are two approaches to the simulation of the time series $\vec{B}$. The first approach is that the “real” BISCR is calculated based on real data for the atmospheric pressure, wind speed modulus, air temperature and relative humidity. After that, the values obtained are used to evaluate parameters of a stochastic model of the BISCR. In the second approach, parameters of the stochastic model of the joint meteorological time series are estimated from observation data and then “artificial” BISCR series are constructed on the simulated trajectories of the joint time series. Each of these approaches has its advantages and disadvantages. Models of the first type are numerically implemented faster than models of the second type (since when using the first approach, it is necessary to simulate the scalar time series, and when using the second – the vector time series). However, models of the second type allow us to

study properties of the BISCR as functions of one or several meteorological parameters. In this paper, we use the second approach to study the properties of the BISCR time series.

The following assumptions will be used:

1. The meteorological time series, as well as the time series of the BISCR, are assumed to be periodically correlated random processes. On the one hand, this assumption allows one to take into account the daily variation of real meteorological processes, and on the other hand - to reduce the simulation complexity as compared to the case when non-stationary (both by one-dimensional distributions and by correlations) time series are simulated.

2. Since the daily variation of the atmospheric pressure time series is mild, the atmospheric pressure $P_i^j,\ i = \overline{1,8},\ j = \overline{1,N}$ is assumed to be constant and equal to the sample average pressure calculated over the considered N days interval. Due to this assumption, instead of the joint series of the four meteorological parameters, it is possible to simulate the time series $\vec{C} = \left(\overrightarrow{C^1}, \overrightarrow{C^2}, ..., \overrightarrow{C^N}\right)$ of the three parameters, where

$$\overrightarrow{C^j} = \left(T_1^j, ..., T_8^j, R_1^j, ..., R_8^j, V_1^j, ..., V_8^j\right)^{\mathrm{T}},\quad j = \overline{1,N},$$

   T_i^j, R_i^j, V_i^j are the values of air temperature, relative humidity and the wind speed modulus at the measurement number i per a day number j ($i = \overline{1,8}, j = \overline{1,N}$).

3. It is also assumed that the one-dimensional distribution of T_i^j is a mixture of the two Gaussian distributions. In this paper, the parameters of the mixtures for all $i = \overline{1,8}, j = \overline{1,N}$ were chosen using the algorithm, proposed in [13]. It is shown in [10] that the use of such a method for approximating the sample one-dimensional density of air temperature distribution allows one to simulate the time series of this weather element with a high accuracy.

   In [11], the sample one-dimensional distributions of relative humidity, depending on the weather station under consideration, were approximated with mixtures of beta distributions, mixtures of truncated Gaussian or gamma-distributions. In this paper, for unification, we use piecewise-linear approximation of the sample distribution function of a relative humidity R_i^j.

   To approximate the one-dimensional distribution densities of the wind speed modulus, the Weibull distribution is often used (see, for example, [2, 8]). However, the numerical experiments have shown that approximation with a mixture of the two gamma-distributions with the mixture parameters determined according to the algorithm, presented in [14], allows one to reproduce more precisely the sample coefficients of asymmetry and kurtosis (in comparison with the approximation with the Weibull distribution). The usage of the mixtures of two gamma-distributions also minimizes the deviation of the approximating

density from the sample one in the sense of the Pearson functional. Therefore, in this paper, we assume that V_i^j has one-dimensional distributions in the form of a mixture of two gamma-distributions.

4. For the simulation of the joint time-series $\vec{C}$, the sample correlation matrix R was used. Analysis of real data shows that for all meteorological stations and time intervals considered, the amplitudes of diurnal oscillations of the auto- and cross-correlation functions of meteorological elements are significant. For estimating the sample correlation coefficients, the biased estimator was used [15].

   For the simulation of the joint time series $\vec{C}$ with given one-dimensional distributions of $T_i^j,\ R_i^j,\ V_i^j,\ i = \overline{1,8}, j = \overline{1,N}$ and a given correlation matrix R, the method of inverse distribution function was used. The simulation of an auxillary standard Gaussian sequence was done using the Cholesky decomposition of its correlation matrix [16].

3 Numerical experiments

Any stochastic model has to be verified before one starts to use the simulated trajectories to study properties of a simulated process. For the model verification, it is necessary to compare the simulated and the real data based on estimations of such characteristics, which, on the one hand, are reliably estimated by real data, and on the other hand are not input parameters of the model. A number of examples of such characteristics are presented below. Although all examples in this paper are given only for the stations in the cities of Tomsk (West Siberia, Russia; subarctic cyclonic climate) and Pogranichniy (Russian Far East; arctic climate), all the conclusions are valid for all the considered weather stations situated in different climatic zones. For the verification, the data collected in 1966-2016 were used.

It should be noted that hereafter 10^6 simulated trajectories for estimations were used . To denote the estimations based on the real and simulated data, abbreviations RD and MD are used, respectively. Hereinafter, σ is the statistical estimate of the standard deviation of the characteristic under consideration when estimating with real data.

The first characteristic to be compared was the average number AN of days over the considered time interval with the average daily BISCR belonging to the given interval. The intervals were chosen according to distinguished levels of discomfort, for example, the BISCR from 10 to 8 points is the level of comfort and the BISCR from 4 to 0 points is the level of non-compensable discomfort [4]. Table 1 shows the corresponding estimations, obtained based on of the real and simulated data. The estimations based on real and simulated data agree reasonably well.

The next characteristic used for the comparison of the real and simulated BISCR time series is the average number $AD\,(l)$ of days over the interval considered, in which the daily minimum BISCR is below the specified level l. The numerical analysis shows that for most of the considered time-intervals, the simulated data

based estimations belong to the corresponding the real data confidence intervals $(AD(l) - \sigma, AD(l) + \sigma)$, see, for example, Table 2.

One more characteristic that can be used to investigate how much the stochastic model of the joint time-series $\vec{C}$ fits for studying the properties of the BISCR is the portion $s(v, n)$ of periods of n measurements with the BISCR not exceeding v among all the periods of measurements. Tables 3 and 4 present examples of estimations of $s(v, n)$ from real and simulated data. The numerical analysis shows that for almost all the considered weather stations, time intervals, levels v and lengths n, the deviations of estimations based on the simulated data from the corresponding estimations from real data do not exceed 3σ.

We have also compared the probabilities $P_{24}(\Delta)$ of the BISCR change greater than by Δ in 24 hours, estimated on real and simulated data. Examples of the estimations of the $P_{24}(\Delta)$ are shown in Table 5. In these examples, the probabilities, estimated based on of the simulated data, belong to the corresponding intervals $(P_{24}(\Delta) - 3\sigma, P_{24}(\Delta) + 3\sigma)$, calculated withe the use of real data, but for other weather stations and time-intervals, in, approximately, 25% of cases (mostly for $\Delta < 0.8$) this is not true. The reason why the model proposed unsatisfactorily reproduces this characteristic of the joint time-series $\vec{C}$, is unclear and calls for further investigations.

The results of the numerical analysis show that the trajectories of the model proposed are close in their statistical properties to the real time series of the BISCR. Therefore, it is possible to use the model in question to study the extreme weather events that are characterized by unfavorable values of the BISCR and to study the dependence of the BISCR properties on the climate change.

Table 1: The Average Number AN of Days on the Considered Time Interval with the Average Daily BISCR Belonging to the Given Interval

Interval	Pogranichniy, July 1–15		Tomsk, January 16–30	
	RD $AN \pm \sigma$	MD	RD $AN \pm \sigma$	MD
$[10; 8]$	0.065 ± 0.048	0.107	0.000 ± 0.000	0.000
$(8; 7]$	13.065 ± 0.241	13.245	0.000 ± 0.000	0.004
$(7; 6]$	1.870 ± 0.234	1.645	0.900 ± 0.028	1.094
$(6; 5]$	0.000 ± 0.000	0.000	8.460 ± 0.531	8.500
$(5; 4]$	0.000 ± 0.000	0.000	5.440 ± 0.533	5.159
$[4; 0]$	0.000 ± 0.000	0.000	0.200 ± 0.001	0.244

Conclusions

In this paper it is shown that the proposed stochastic model of the joint time series of the wind speed modulus, air temperature and relative humidity may be used for

Table 2: The Average Number $AD\,(l)$ of Days on the Considered Interval, in Which the Daily Minimum BISCR is Below the Level l

l	Pogranichniy, January 1–15		Tomsk. July 1–15	
	RD $AD\,(l) \pm \sigma$	MD	RD $AD\,(l) \pm \sigma$	MD
10.0	15.000 ± 0.000	15.000	15.000 ± 0.000	15.000
9.5	15.000 ± 0.000	15.000	14.920 ± 0.032	14.948
9.0	15.000 ± 0.004	14.999	11.580 ± 0.281	11.425
8.5	15.000 ± 0.010	14.996	6.600 ± 0.357	5.707
8.0	14.978 ± 0.023	14.981	2.680 ± 0.255	2.170
7.5	14.913 ± 0.066	14.851	0.620 ± 0.127	0.562
7.0	13.783 ± 0.204	13.707	0.120 ± 0.042	0.073
6.5	10.022 ± 0.386	9.676	0.000 ± 0.006	0.006
6.0	3.957 ± 0.348	3.824	0.000 ± 0.002	0.000
5.5	0.826 ± 0.148	0.6690	0.000 ± 0.001	0.000
5.0	0.022 ± 0.036	0.049	0.000 ± 0.000	0.000
4.5	0.000 ± 0.006	0.002	0.000 ± 0.000	0.000
4.0	0.000 ± 0.001	0.000	0.000 ± 0.000	0.000

Table 3: Probabilities $s\,(v, n)$. Tomsk. January, 1-15

n	$v = 7$		$v = 6$	
	RD $s\,(v, n) \pm 3\sigma$	MD	RD $s\,(v, n) \pm 3\sigma$	MD
2	0.995 ± 0.002	0.999	0.913 ± 0.045	0.899
4	0.991 ± 0.003	0.999	0.877 ± 0.059	0.855
6	0.987 ± 0.004	0.999	0.849 ± 0.069	0.820
8	0.983 ± 0.005	0.998	0.824 ± 0.078	0.790
10	0.978 ± 0.007	0.998	0.804 ± 0.086	0.762
12	0.973 ± 0.008	0.998	0.785 ± 0.093	0.736
14	0.969 ± 0.009	0.997	0.769 ± 0.010	0.712

studying the probabilistic properties of the high resolution time series of the bioclimatic index of severity of climatic regime on the time intervals on which a seasonal variation has not a significant influence on the real meteorological processes. In the future, similar studies will be conducted on longer time intervals. Since the characteristics of the atmospheric pressure change significantly over long time intervals, instead of the model of a periodically correlated complex "air temperature - relative humidity - wind speed modulus", a model of the non-stationary joint time series "temperature - relative humidity - wind speed modulus - atmospheric pressure" will be used. This will allow us to take into account both the daily and the seasonal variations of real

Table 4: Probabilities $s\,(v,n)$. Tomsk. January, 1-15

n	$v=5$		$v=4$	
	RD $s\,(v,n)\pm 3\sigma$	MD	RD $s\,(v,n)\pm 3\sigma$	MD
2	0.321 ± 0.084	0.293	0.015 ± 0.018	0.019
4	0.254 ± 0.078	0.223	0.007 ± 0.013	0.010
6	0.213 ± 0.073	0.181	0.003 ± 0.010	0.006
8	0.185 ± 0.068	0.150	0.002 ± 0.008	0.004
10	0.164 ± 0.064	0.126	0.002 ± 0.007	0.003
12	0.146 ± 0.060	0.106	0.001 ± 0.005	0.002
14	0.131 ± 0.057	0.090	0.001 ± 0.004	0.001

Table 5: Probabilities $P_{24}\,(\Delta)$ of the BISCR change more than by Δ in 24 hours

Δ	Tomsk, July 16-30		Pogranichniy, July 16-30	
	RD $P_{24}\,(\Delta)\pm 3\sigma$	MD	RD $P_{24}\,(\Delta)\pm 3\sigma$	MD
0.1	1.0000 ± 0.000	1.000	1.000 ± 0.000	1.000
0.5	0.980 ± 0.012	0.986	0.887 ± 0.026	0.935
1.0	0.832 ± 0.046	0.858	0.454 ± 0.075	0.522
1.5	0.524 ± 0.072	0.591	0.173 ± 0.059	0.189
2.0	0.224 ± 0.061	0.268	0.049 ± 0.031	0.056

meteorological processes.

Acknowledgements

This work was partly financially supported by the Russian Foundation for Basic Research (grant No 18-01-00149-a), by the Russian Foundation for Basic Research and the government of the Novosibirsk region according to the research project No 19-41-543001-r-mol-a.

References

[1] Anderson G.B., Bell M.L., Peng R.D. (2013). Methods to Calculate the Heat Index as an Exposure Metric in Environmental Health Research. *Environmental Health Perspectives*. Vol. **121**, No 10, pp. 1111-1119.

[2] Balouktsis A., Tsanakas D., Vachtsevanos G. (1986). Stochastic Simulation of Hourly and Daily Average Wind Speed Sequences. *Wind Engineering*. Vol. **10**, No 1, pp. 1-11.

[3] Bylaw No 107-r of the Government of the Russian Federarion of 18.01.1992

[4] Belkin V.S., Dyurgerov M.B., Finaev A.F., Soroko S.I. (2016). Bioclimatic evaluation of the human discomfort level for several Antarctic regions. *Human Physiology*. Vol. **42**, No 2, pp. 119-127.

[5] Belkin V.S., Poltorak G.I. (1983). Some biomedical aspects of studying mountain regions of Tajikistan. *Congress of the Geographical Society in Dushanbe*. pp. 19-21. (in Russian)

[6] Belkin V.S., Sokolov L.N., Finaev A.F. (1989). Special features of the bioclimatic zoning of mountain areas. *Abstracts of the All-Union meeting on human bioclimatology*. pp. 15-16. (in Russian)

[7] Finaev A.F. (2004). Assessment of the impact of climatic conditions on human labor.*Archeology and Paleoecology of Eurasia*. Ed. by A.P. Derevyanko. SB RAS, Novosibirsk. pp. 359-376. (in Russian)

[8] Gevorkyan M.N., Demidova A.V., Sobolewski R.A., Zaryadov I.S., Korolkova A.V., Kulyabov D.S., Sevastianov L.A. (2017). Approaches to Stochastic Modeling of Wind Turbines. https://arxiv.org/pdf/1711.03589.pdf

[9] Hill L.E., Angus T.C., Newbold E.M. (1958). Further experimental observations to determine the relations between kata cooling powers and atmospheric conditions. *J. Ind. Hyg.* Vol. **10**, pp. 391–407.

[10] Kargapolova N.A. (2018). Monte Carlo Simulation of Non-stationary Air Temperature Time-Series. *Proc. of 8th Int. Conf. on Simulation and Modeling Methodologies, Technologies and Applications "SIMULTECH-2018"*. pp. 323-329.

[11] Kargapolova N.A., Khlebnikova E.I., Ogorodnikov V.A. (2018). Monte Carlo simulation of the joint non-Gaussian periodically correlated time-series of air temperature and relative humidity.*Statistical papers*. Vol. **59**, pp. 1471-1481.

[12] Kobisheva N.V., Stadnik V.V., Klueva M.V., Pigoltsina G.B., Akentieva E.M., Galuk L.P., Razova E.N., Semenov U.A. (2008). *Guidance on specialized climatological service of the economy*. Asterion, St.Petersburg. (in Russian)

[13] Marchenko A.S., Minakova L.A. (1980). Probabilistic model of air temperature time-series.*Meteorology and Hydrology*. No. 9, pp. 39-47. (in Russian)

[14] Marchenko A.S., Syomochkin A.G. (1982). Models of unidimensional and joint distributions of non-negative random values. *Meteorology and Hydrology*. No 3, pp. 50-56. (in Russian)

[15] Marple S.L. Jr. (1987). *Digital Spectral Analysis with Applications*. Prentice-Hall, Englewood Cliffs, NJ.

[16] Ogorodnikov V.A., Prigarin S.M. (1996). *Numerical Modelling of Random Processes and Fields: Algorithms and Applications*. VSP,Utrecht.

[17] Osczevski R., Bluestein M. (2005). The New Wind Chill Equivalent Temperature Chart. *Bulletin of the American Meteorological Society*. Vol. **86**, pp. 1453-1458.

Approximate numerical stochastic spectral model of a periodically correlated process

Alisa M. Medvyatskaya[1] and Vasily A. Ogorodnikov[1],[2]
[1] *Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
e-mail: medvyatskaya@mail.ru, ova@osmf.sscc.ru

Abstract

This research is devoted to the development of an approximate numerical stochastic modeling based on real data of a periodically correlated Gaussian process based on a spectral representation in which the Fourier coefficients form a stationary vector Gaussian process.

Keywords: spectral model, vector process, matrix correlation function.

Introduction

In [5], a spectral model of a scalar periodically correlated random process was considered, based on the representation of

$$\xi(t) = \sum_{p \in Z} \eta_p(t) \exp\left(i \frac{2\pi}{T} pt\right), \tag{1}$$

and $\eta_p(t)$ - are the components of the vector infinite-dimensional stationary random process with zero means and the matrix correlation function $K_{pq}(t-s)$, $p, q \in Z$, T- is an arbitrary deterministic value having the dimension of time. This process is a periodically correlated process with a period T [4] that has such properties

$$E\eta(t) = E\eta(t+T), \quad D\eta(t) = D\eta(t+T),$$
$$E(\eta(t) - E\eta(t))(\eta(s) - E\eta(s)) = R(t,s) = R(t+T, s+T).$$

In our research it will be shown that covariance function of this process is

$$R(t,s) = \sum_{p \in Z} K_{pq}(t-s) \exp\left(i \frac{2\pi}{T}(pt - qs)\right)$$

and this function covers the entire class of covariance functions of an arbitrary periodically correlated process. In this study, this representation is used to construct an approximate stochastic model of a periodically correlated random process, in which an infinite-dimensional random vector $\eta(t)$ is replaced by a finite-dimensional dimension n. Here we note that the algorithms for constructing periodically correlated processes based on other approaches were considered in [6], [8], [7] based on Markov chains with a periodically changing transition probability matrix, based on Poisson flows with a periodic intensity, based on spectral representation with periodically varying

spectral density. The most common approach to modeling periodically correlated processes is an approach based on modeling vector Gaussian stationary processes of a discrete argument with a constant time step, which also allows you to simulate non-Gaussian periodically correlated sequences based on the method of inverse distribution functions [3], [4]. Models based on the spectral representation make it possible to model processes with an uneven time step [30], [31] without additional computational costs.

1 Construction of a spectral model of periodically correlated processes with using real data

We consider a periodically correlated random process $\xi(t)$ of continuous time in the interval $[0, \infty)$ with a period of correlation T. We divide the time axis into intervals $[(k-1)T+1,\ kT+1]$, $k = 1, 2, \ldots$. Consider the representation (1) in the form

$$\xi(t) = \sum_{p=0}^{n} \left(\eta_p(k) \cos\left(\frac{2\pi}{T} pt\right) + \zeta_p(k) \sin\left(\frac{2\pi}{T} pt\right) \right), \tag{2}$$

where $t \in [(k-1)T+1,\ kT+1]$, $\eta_p(k)$ and $\zeta_p(k)$ - are components of stationary vector Gaussian random processes with mean $\mu_{\eta p}(k)$ and $\nu_{\zeta p}(k)$, variances $\sigma^2_{\eta p}(k)$ and $\sigma^2_{\zeta p}(k)$ and a joint matrix correlation function

$$\Phi(\tau) = \begin{pmatrix} \Phi_{\eta\eta}(\tau) & \Phi_{\eta\zeta}(\tau) \\ \Phi_{\zeta\eta}(\tau) & \Phi_{\zeta\zeta}(\tau) \end{pmatrix},$$
$$\tau = k - p, \quad k, p = 0, 1, \ldots$$

Algorithm for numerical simulation of the process (2) and methods for estimating the parameters of the model based on real data are considered in this study. In meteorological tasks, in order to choose one or another approximation, for example, stationarity or periodic correlation, time intervals are considered at relatively short time intervals, most often one month long or less. Time series for the same moments of time within a year, but corresponding to different years of observation, are approximately considered independent and are analogous to the trajectories of the process $\xi(t)$. Lets represent l-th trajectory of a periodically correlated process $\tilde{\xi}_l(t)$, $l = 1, 2, \ldots$ with a discrete argument and a period of correlation equal to $T = m$ in the form of

$$\tilde{\xi}_l(t_1), \ldots, \tilde{\xi}_l(t_m), \tilde{\xi}_l(t_{m+1}), \ldots, \tilde{\xi}_l(t_{2m}), \ldots \tilde{\xi}_l(t_{(k-1)m+1}), \ldots, \tilde{\xi}_l(t_{km}), \ldots.$$

We divide the range of values of its argument into intervals $[(k-1)T+1,\ kT+1]$, $k = 1, 2, \ldots$. Let $m+1$ - the number of measurements in each of these intervals. In each k-th interval, process values are given in a discrete sequence of points.

$$(k-1)T+1 = t_{(k-1)m+1} < t_{(k-1)m+2} < \cdots < t_{(k-1)m+m} < t_{(k-1)m+m+1} = kT+1.$$

The process values $\tilde{\xi}_l(t) = \tilde{\xi}_{l,k}(t)$ in these points are denoted as

$$\tilde{\xi}_{l,k}(t_{(k-1)m+1}), \ldots, \tilde{\xi}_{l,k}(t_{km+1}),$$

where k - is interval number. Interpolate $\tilde{\xi}_{l,k}(t)$ in the interval $[(k-1)T+1,\ \ kT+1]$, $k = 1, 2, \ldots$, by its values $\tilde{\xi}_{l,k}(t_{(k-1)m+1}), \ldots, \tilde{\xi}_{l,k}(t_{km+1})$ at these points. In each k-th interval $[(k-1)T+1,\ \ kT+1]$ we represent l-th trajectory $\tilde{\xi}_l(t),\ l = 0, 1, \ldots$ process $\xi(t)$ as a Fourier series

$$\tilde{\xi}_l(t) = \tilde{\xi}_{l,k}(t) = \sum_{p=0}^{n} \left(\tilde{\eta}_{p,l}(k) \cos\left(\frac{2\pi}{T} pt\right) + \tilde{\zeta}_{p,l}(k) \sin\left(\frac{2\pi}{T} pt\right)\right) \quad , \tag{3}$$

where $\tilde{\eta}_{p,l}(k)$ and $\tilde{\zeta}_{p,l}(k)$ are the Fourier coefficients and these coefficients are determined by the expressions

$$\tilde{\eta}_{p,l}(k) = \frac{2}{T} \int\limits_{-T/2}^{T/2} \tilde{\xi}_{l,k}(t) \cos(\frac{2\pi}{T} pt) dt, \quad t \in [(k-1)T+1,\ \ kT+1],$$

$$\tilde{\zeta}_{p,l}(k) = \frac{2}{T} \int\limits_{-T/2}^{T/2} \tilde{\xi}_{l,k}(t) \sin(\frac{2\pi}{T} pt) dt, \quad t \in [(k-1)T+1,\ \ kT+1], \tag{4}$$

$$p = 1, 2, \ldots, n,$$

$$\tilde{\eta}_{0,l}(k) = \frac{1}{T} \int\limits_{-T/2}^{T/2} \tilde{\xi}_{l,k}(t) dt, \quad t \in [(k-1)T+1,\ \ kT+1], \quad \tilde{\zeta}_{0,l}(k) = 0.$$

and for each l form scalar sequences $\tilde{\eta}_{p,l}(k)$ and $\tilde{\zeta}_{p,l}(k)$, $p = 0, 1, \ldots, n$, $k = 0, 1, \ldots, K$. We represent these sequences as a single vector sequence

$$\vec{\tilde{\phi}}_{l,0}, \vec{\tilde{\phi}}_{l,1}, \ldots, \vec{\tilde{\phi}}_{l,K}, \tag{5}$$

where

$$\vec{\tilde{\phi}}_{l,k} = (\tilde{\eta}_l^T(k), \tilde{\zeta}_l^T(k))^T = (\tilde{\eta}_{0,l}(k), \tilde{\eta}_{1,l}(k), \ldots, \tilde{\eta}_{n,l}(k), \tilde{\zeta}_{1,l}(k), \ldots, \tilde{\zeta}_{n,l}(k))^T \quad - \tag{6}$$

is vector formed from the Fourier coefficients for the l-th trajectory and interval $[(k-1)T+1,\ \ kT+1]$. Here the number of components for the sub-vectors $\tilde{\eta}_l^T(k)$ and $\tilde{\zeta}_l^T(k)$ of the vector $\vec{\tilde{\phi}}_{l,k}$ is equal to $n+1$ and n while the numbering of the second sub-vector begins with $p = 1$. The sequence (5) forms a sample of a stationary sequence $\vec{\tilde{\phi}}_{l,0}, \vec{\tilde{\phi}}_{l,1}, \ldots$ of vectors formed from the Fourier coefficients $\tilde{\eta}_{p,l}(k)$ and $\tilde{\zeta}_{p,l}(k)$. The stationarity of this sequence directly follows from the periodic correlation of the process $\xi(t)$.

It is assumed that the period of correlation is known in advance. For example, when modeling meteorological series, this period is usually determined by a daily or annual cycle. If the period of correlation is determined by other factors, then special additional studies are needed to determine it [5], [9],. The algorithm for simulating a Gaussian periodically correlated sequence based on the representation (2) and relations (3)-(5) based on real data is reduced to the following transformations:
Algorithm

1. The original time interval is divided into sub-intervals $[(k-1)T+1,\ kT+1]$, $k=\ 1,\ \ldots, K$.

2. The values of the series k-th in the subinterval $\tilde{\xi}_{l,k}(t_{(k-1)m+1}), \ldots, \tilde{\xi}_{l,k}(t_{km+1})$ are interpolated to each point of this interval. As a result, we obtained the functions $\tilde{\xi}_{l,k}(t),\ t \in\ [(k-1)T+1,\ kT+1]$.

3. Integrals (4) are calculated for each value $k=\ 1,\ \ldots, K,\ l=1,2,\ldots,L$, $p=0,1,\ldots,n$

4. For each $p=0,1,\ldots,n\ \ k=1,\ldots,K,\ l=1,2,\ldots,L$ vector sequences $\vec{\phi}_{l,0}, \vec{\phi}_{l,1}, \ldots, \vec{\phi}_{l,K}$ (5) are formed from the obtained values $\tilde{\eta}_{p,l}(k)$ and $\tilde{\zeta}_{p,l}(k)$, forming a sample of vector sequences length K in which the dimension of the vectors is $2n+1$.

5. For this sample, the mean values $\vec{\mu}=(\mu_1,\ldots,\mu_{2n-1})^T$, variances $\vec{\sigma}^2=(\sigma_1^2,\ldots,\sigma_{2n-1}^2)^T$ and matrix correlation function $\Phi(\tau)$ of the vector sequence of Fourier coefficients are estimated.

6. Next, a Gaussian stationary vector sequence $\vec{\phi}_{l,0}, \vec{\phi}_{l,1}, \ldots, \vec{\phi}_{l,K}$ is simulated with zero mean, unit variance and correlation matrix $\Phi(\tau)$.

7. Based on the obtained vector sequence, a stationary vector sequence of model Fourier coefficients is constructed by multiplying the obtained components of the Gaussian vectors $\vec{\phi}_{l,0}, \vec{\phi}_{l,1}, \ldots, \vec{\phi}_{l,K}$ by the corresponding sample standard deviations and additions with the corresponding sample means $\mu_i,\ i=1,\ldots,2n+1$.

8. The final periodically correlated process is constructed using the relation (3) for $t \in\ [(k-1)T+1,\ kT],\ k=1,\ldots,K$.

2 Numerical experiments

As an example, in this paper, the considered algorithm is used to simulate periodically correlated time series of air temperature according to four-time observations of the temperature at the Sverdlovsk meteorological station in the month of May for the period from 1936 to 1984 years. The period of correlation in this case is one day (or four periods of observation $T=m=4$). In this research, the values of a real time

series $\tilde{\xi}(t_1), \tilde{\xi}(t_2), \ldots, \tilde{\xi}(t_M)$ correspond to conditional time points $t_1, t_2, \ldots, t_M$, $t_i = i$ with a step $\Delta t = t_{i+1} - t_i = 1$. Here $M = K \times T = 31 \times 4$ (K- number of days in May, $T = t_{i+m} - t_i = 4$ - period of correlation). The original time interval is divided into subintervals $[(k-1)T + 1,\ kT + 1]$, $k = 1, \ldots, K$. The interpolation is a piecewise linear interpolation from the Algorithm point 2. In each k-th interval $[(k-1)T + 1,\ kT + 1]$ the integrals (4) are calculated by the Simpson method. And further, Algorithm points 5-7 are implemented.

The Gaussian vector sequence $\vec{\phi}_{l,0}, \vec{\phi}_{l,1}, \ldots, \vec{\phi}_{l,K}$ is simulated based on the spectral decomposition of the matrix $\Phi(\tau)$. In the calculations, n was chosen equal to 9. $L = 10000$ of process trajectories were used in the calculations. The matrix correlation functions $R(\tau)$ of the model $\xi(t)$ and real $\tilde{\xi}_l(t)$, $i = 1, \ldots, L$ processes at the points $t_1, t_2, \ldots, t_M$, $t_i = i$ were calculated along these trajectories. We note that a periodically correlated scalar process defined at equally spaced points is equivalent to a vector stationary process in which the dimension of the vectors is equal to the correlation period [5]. Therefore the correlation function of a periodically correlated scalar process is equivalent to the matrix correlation function $R(\tau)$ of a vector stationary process. These functions for model and real processes are shown in Table1.

Table 1.

Model and real matrix correlation functions of the process

	Modeldata				*Realdata*			
t_i	$\tau = 0$				$\tau = 0$			
1	1.000	0.848	0.718	0.701	1.000	0.834	0.709	0.694
2	0.848	1.000	0.841	0.704	0.834	1.000	0.837	0.692
3	0.718	0.841	1.000	0.904	0.709	0.837	1.000	0.897
4	0.701	0.704	0.904	1.000	0.694	0.692	0.897	1.000
t_i	$\tau = 1$				$\tau = 1$			
1	0.669	0.517	0.426	0.435	0.678	0.516	0.437	0.460
2	0.640	0.659	0.510	0.411	0.638	0.665	0.523	0.422
3	0.791	0.772	0.645	0.543	0.792	0.781	0.662	0.560
4	0.892	0.785	0.701	0.666	0.894	0.784	0.709	0.677
t_i	$\tau = 2$				$\tau = 2$			
1	0.423	0.293	0.233	0.276	0.459	0.327	0.283	0.329
2	0.359	0.379	0.286	0.241	0.380	0.416	0.333	0.281
3	0.480	0.459	0.380	0.327	0.505	0.493	0.422	0.365
4	0.615	0.494	0.415	0.411	0.634	0.514	0.446	0.445
t_i	$\tau = 3$				$\tau = 3$			
1	0.287	0.183	0.147	0.170	0.335	0.219	0.200	0.256
2	0.209	0.252	0.184	0.130	0.249	0.296	0.251	0.213
3	0.296	0.293	0.208	0.160	0.325	0.331	0.280	0.241
4	0.384	0.282	0.219	0.232	0.424	0.321	0.281	0.303

The difference in the values of the model and real correlation matrices is determined by such factors as, for example, the accuracy of the calculation of the integrals, the limited number of harmonics used. If the estimation of the matrix correlation function is carried out at points different from $t_1, t_2, \ldots, t_M$, $t_i = i$,other than, the corresponding values of the model matrix are determined by the method of interpolation of the initial vector sequence in the intervals $[(k-1)T+1,\ kT+1]$, $k = 1, \ldots, K$.

Conclusion

In conclusion, we note that by increasing the number of harmonics and the accuracy of calculating the integrals, we can significantly increase the accuracy of modeling. The accuracy of the simulation is checked by the magnitude of the proximity of the sample and model matrix correlation functions, which are estimated at the time points of real observations. The real correlation function of the process is not an input information to the model, but the matrix correlation function of vectors from the Fourier coefficients that significantly depends on the sample size, so the sample size implicitly affects the accuracy of the simulation. The model reproduces the real process, which is initially set on a regular grid in arbitrary points of the considered time domain and everywhere, except for the moments in which the initial process is specified, is determined to the interpolation method. Thus it is of further interest to study the dependence of the properties of the correlation function of the simulated process on the choice of interpolation method. Probably the considered approach can be generalized to the case of a nonstationary process of a more general form. It should also be noted that interpolation methods can also be applied to the methods considered in [3], [5], however, the difference is that in this investigation interpolation is applied to the source data within intervals [(k-1)T+1, kT+1], k=1,...,K, and in methods based in the simulation of stationary vector processes, interpolation should be applied to the modeling process.

Acknowledgements

The reported study was funded by RFBR (grant 18-01-00149), RFBR and the government of the Novosibirsk region according to the research project № 19-41-543001-r-mol-a

References

[1] Bokov V. N., Lopatukhin L. I., Mikulinskaya S. M., Rozhkov V. A., Rumyantseva S. A. (1995) On The Interannual Variability Of The Disturbances. Saint Petersburg: Gidrometeoizdat, pp. 446 – 454. (in Russian)

[2] Derenok K.V., Ogorodnikov V.A. (2008). Numerical simulation of significant long-term decreases in air temperature // Russ. J. Num. Anal. Math. Modelling, Vol. **23**, No 3. pp. 223 – 277.

[3] Dragan Y. P. , Rozhkov V. A. ,Yavorsky I. N. (1987). The Methods of Probabilistic Analysis of Oceanological Rhythmics, L: Gidrometeoizdat, pp. 320 (in Russian)

[4] Kargapolova N.A., Ogorodnikov V.A.(2012) Inhomogeneous Markov chains with periodic matrices of transition probabilities and their application to simulation of meteorological processes // Russ. J. Num. Anal. Math. Modelling, Vol. **27**, No 3.pp. 213 – 228.

[5] Medvyatskaya A.M. (2016) Spectral model of a class of periodically correlated processes // works of the conference of young scientists The Institute of Computational Mathematics and Mathematical Geophysics SB RAS (ICM&MG SB RAS). pp. 36-41. (in Russian)

[6] Mikhailov G.A., Voytishek A.V. (2006) Numerical Statistical Modeling // Moscow: Publishing center Academy, pp. 368 (in Russian)

[7] Ogorodnikov V.A., Sereseva O.V., Kargapolova N.A. (2016) Stochastic models of piecewise-constant and piecewise-linear non-Gaussian processes based on Poisson flow // Russ. J. Num. Anal. Math. Modelling, Vol. **3**1, No 3. pp. 179-185.

[8] Prigarin S.M.(2005). Numerical Modeling Of Random Processes And Fields, ICM&MG Publisher, Novosibirsk, pp. 259. (in Russian)

[9] Rozhkov V.A., Trapeznikov Yu.A.(1990) Probabilistic models of oceanological processes. – L: Gidrometeoizdat, pp. 272. (in Russian)

Modeling of dispersion in a fractal porous medium

Olga N. Soboleva[1,2,3]
[1] *Institute of Computational Mathematics and Mathematical Geophysics*
[2] *Novosibirsk State University,*
[3] *Novosibirsk State Technical University, Novosibirsk, Russia*
e-mail: olgasob@gmail.com

Abstract

The effective coefficients for equations of the convective diffusion are obtained. The correlated fields of conductivity and porosity are approximated by the Kolmogorov multiplicative continuous cascades with a lognormal probability distribution. The theoretical results for the incompressible flow are compared with the results from direct 3D numerical simulations.

Keywords: effective coefficients, convective diffusion, lognormal probability distribution, multiplicative continuous cascades.

Introduction

In natural conditions, as a rule, the spatial geometry of small-scale heterogeneities is not exactly known, and the irregularity of conductivity and porosity abruptly increases as the scale of measurements decreases. Since generally porosity and conductivity vary in an irregular manner, it is customary to regard them as random space functions characterized by the joint probability distribution functions and they are taken into account with the help of the effective parameters. Many natural media are "scale regular" in the sense that they can be described by multifractals and hierarchical cascade models. In this paper, by the method of subgrid modeling, we obtain formulas of effective coefficients for for equations of the convective diffusion. The effective coefficients depend not only on means and variances of the parameters, but also on the correlation between the conductivity and porosity. The theoretical results are verified with the help of direct 3D numerical simulations.

1 Governing equations and model of the medium

At low Reynolds numbers, the filtration velocity $\mathbf{v}$ and the pressure are related by Darcy law $\mathbf{v} = -\sigma(\mathbf{x})\nabla p$, where a random function of the spatial coordinates $\sigma(\mathbf{x})$ is a conductivity coefficient. The condition of incompressibility yields the equation

$$\nabla[\sigma(\mathbf{x})\nabla p(\mathbf{x})] = 0, \tag{1}$$

where $\mathbf{x}$ is the three-dimensional vector of spatial coordinates, $p(\mathbf{x})$ is the pressure. At the initial time a colored liquid flow-in into a volume filled with a pure liquid. The interface is labelled with passive particles, which are moved by a stationary velocity field. Since both liquids have the same physical parameters, their filtration velocities

satisfies the Darcy equation (1). The movement of the labelled particles is described by the equation [1]

$$m(\mathbf{x}_i)\frac{d\mathbf{x}_i(t)}{dt} = -\sigma(\mathbf{x}_i)\nabla p, \quad \mathbf{x}_i(0) = \mathbf{x}_{i0}, : \tag{2}$$

where $i = 1, \ldots, N$ is the number of a particle. The pressure p is defined by equation (1) and gradient is calculated at the point $\mathbf{x}_i$, $m(\mathbf{x})$ is the porosity coefficient.

Let the conductivity field be known. This means that it is measured at each point as the fluid is pumped through a sample of small size. In an experiment, one measures a conductivity fields within the accuracy of a minimal scale l_0. A random function of spatial coordinates $\sigma(\mathbf{x})$ is considered as limit of conductivity, as, we have $l_0 \to 0$, $\sigma(\mathbf{x})_{l_0} \to \sigma(\mathbf{x})$. Let the conductivity field satisfies $\sigma(\mathbf{x}) = \sigma(\mathbf{x})_{l_0}$. To pass to a coarser grid l_1, one can smooth the resultant field $\sigma(\mathbf{x})_{l_0}$ using the scale $l_1 > l_0$. The obtained field is not the true conductivity that describes filtration in the interval of scales (l_1, L), where L is the maximum scale of heterogeneities. To find conductivity on a coarser grid one has to repeat the measurements, pumping the fluid through larger sample of size l_1. This procedure is necessary, since the fluctuations of conductivity within the scale interval (l_0, l_1) have correlations with pressure fluctuations induced by them (equation (1)). The search of a transformation law of the effective conductivity, when the scale grid varies, is not so difficult for the self-similar medium. Similar to [3], we consider a dimensionless field ψ equal to the ratio of conductivity smoothed using two different scales (l, l_1). More detail this approach have been described in [2]. The field $\sigma(\mathbf{x})_l$ is $\sigma(\mathbf{x})_{l_0}$ smoothed over scale l, $\psi(\mathbf{x}, l, l_1) = \frac{\sigma(\mathbf{x})_{l_1}}{\sigma(\mathbf{x})_l}$. If $l_1 \to l$ we obtain $\partial \ln \sigma(\mathbf{x},l) / \partial \ln l = \varphi(\mathbf{x},l)$, where function $\varphi(\mathbf{x}, l) = \frac{d\psi(\mathbf{x},l,l\lambda)}{d\lambda}|_{\lambda=1}$ defines the all statistical properties of porous medium [2]. The solution to this equation gives the conductivity as the function of field φ with the given distribution:

$$\sigma(\mathbf{x})_{l_0} = \sigma_0 \exp\left[-\int_{l_0}^{L} \varphi(\mathbf{x}, l)\frac{dl}{l}\right], \tag{3}$$

where σ_0 is constant. The field φ determines statistical properties of the conductivity. According to the limit theorem for sums of independent random variables, if the variance of $\varphi(\mathbf{x}, l)$ is finite, the integral in (3) tends to a field with a normal distribution as the ratio L/l_0 increases. If the variance of $\varphi(\mathbf{x}, l)$ is infinite and there exists a non-degenerate limit of the integral in (3), the integral tends to a field with a stable distribution. In this paper, it is assumed that the conductivity $\sigma(x)_{l_0} \equiv \sigma(\mathbf{x})$ has heterogeneities of the scale l in the interval $l_0 < l < L$ and a correlation function of field $\varphi(\mathbf{x}, l)$ is statistically homogeneous, isotropic

$$\Phi^{\varphi\varphi}(\mathbf{x}, \mathbf{y}, l, l') = \langle \varphi(\mathbf{x}, l)\varphi(\mathbf{y}, l')\rangle - \langle \varphi(\mathbf{x}, l)\rangle \langle \varphi(\mathbf{y}, l')\rangle = \Phi^{\varphi\varphi}(|\mathbf{x} - \mathbf{y}|, l)\, \delta(\ln l - \ln l'). \tag{4}$$

For simplicity, we use the same notation $\Phi^{\varphi\varphi}$ in right-hand side. The angle brackets denote the ensemble averaging. It follows from (4) that the fluctuations of $\varphi(\mathbf{x}, l)$ at different scales do not correlate. This assumption is standard in the scaling models,

see [3], and is due to the fact that the statistical dependence is small if the scales of fluctuations are different. To derive subgrid formulas to calculate effective coefficients, this assumption may be ignored. However, this assumption is important for the numerical simulation of the field $\varphi(\mathbf{x}, l)$. Here we assume that the random field $\varphi(\mathbf{x}, l)$ has the Gaussian distribution. For a scale invariant medium, the following relation holds for any positive K: $\Phi^{\varphi\varphi}(|\mathbf{x}-\mathbf{y}|, l, l') = \Phi^{\varphi\varphi}(K|\mathbf{x}-\mathbf{y}|, Kl, Kl')$.

If for any l the equality $<\sigma(\mathbf{x})_l>=\sigma_0$ is valid (the conservative cascade), then it follows from (3) and (4) that $\Phi_0^{\varphi\varphi} = 2\langle\varphi\rangle$, where $\Phi_0^{\varphi\varphi} = \Phi^{\varphi\varphi}(0, l)$.

The porosity coefficient $m(\mathbf{x})$ is constructed similar to the conductivity coefficient:

$$m(\mathbf{x})_{l_0} = m_0 \exp\left[-\int_{l_0}^{L} \chi(\mathbf{x}, l)\frac{dl}{l}\right]. \tag{5}$$

The function $\chi(\mathbf{x}, l)$ is assumed to have a normal distribution and to be delta correlated in the logarithm of scale. From the physical essence of the porosity follows that the parameters of the porosity field is satisfied: $\Phi_0^{\chi\chi} = 2\langle\chi\rangle$.

The correlation between the porosity and conductivity fields is determined via the correlation of the fields $\varphi(\mathbf{x}, l)$ and $\chi(\mathbf{x}, l)$:

$$\Phi^{\varphi\chi}(\mathbf{x}, \mathbf{x}, l, l') = \langle\varphi(\mathbf{x}, l)\chi(\mathbf{x}, l')\rangle - \langle\varphi(\mathbf{x}, l)\rangle\langle\chi(\mathbf{x}, l')\rangle = \Phi_0^{\varphi\chi}\delta(\ln l - \ln l');$$

2 Effective coefficients

The conductivity function $\sigma(\mathbf{x})$, $m(\mathbf{x})$ are divided into two components with respect to the scale l. The large-scale (ongrid) components $\sigma(\mathbf{x}, l)$, $m(\mathbf{x}, l)$ are obtained, respectively, by statistical averaging over all $\varphi(\mathbf{x}, l_1)$, $\chi(\mathbf{x}, l_1)$ with $l_0 < l_1 < l$, $l - l_0 = dl$, where dl is small. The small-scale (subgrid) components are equal to $\sigma'(\mathbf{x}) = \sigma(\mathbf{x}) - \sigma(\mathbf{x}, l)$, $m'(\mathbf{x}) = m(\mathbf{x}) - m(\mathbf{x}, l)$:

$$\sigma(\mathbf{x}, l) = \sigma_0 \exp\left[-\int_{l}^{L} \varphi(\mathbf{x}, l_1)\frac{dl_1}{l_1}\right]\left\langle \exp\left[-\int_{l_0}^{l} \varphi(\mathbf{x}, l_1)\frac{dl_1}{l_1}\right]\right\rangle, \tag{6}$$

$$\sigma'(\mathbf{x}) = \sigma(\mathbf{x}, l)\left[\frac{\exp\left[-\int_{l_0}^{l} \varphi(\mathbf{x}, l_1))\frac{dl_1}{l_1}\right]}{\left\langle \exp\left[-\int_{l_0}^{l} \varphi(\mathbf{x}, l_1)\frac{dl_1}{l_1}\right]\right\rangle} - 1\right], \ \langle\sigma'(\mathbf{x})\rangle = 0,$$

$$m(\mathbf{x}, l) = m_0 \exp\left[-\int_{l}^{L} \chi(\mathbf{x}, l)\frac{dl}{l}\right]\left\langle \exp\left[-\int_{l_0}^{l} \chi(\mathbf{x}, l)\frac{dl}{l}\right]\right\rangle, \tag{7}$$

$$m'(\mathbf{x}) = m(\mathbf{x}, l)\left[\exp\left[-\int_{l_0}^{l} \frac{dl}{l}\left(\chi(\mathbf{x}, l) - \langle\chi\rangle + \frac{1}{2}\Phi_0^{\chi\chi}\right)\right] - 1\right].$$

We carry out similar partitions for the displacement and the pressure:

$$\mathbf{x}(t) = \mathbf{x}(t,l) + \mathbf{x}'(t)\,,\ p(\mathbf{x}) = p(\mathbf{x},l) + p'(\mathbf{x})\,,\ \langle \mathbf{x}'(t)\rangle = 0,\ \langle p'(\mathbf{x})\rangle = 0,$$

where $\mathbf{x}(t,l)$, $p(\mathbf{x},l)$ are respectively solutions to equations (1), (2) averaged over the small-scale fields σ', m'. Averaging equations (1),(2) over m', σ' with given $m(\mathbf{x},l)$, $\sigma(\mathbf{x},l)$ yields the ongrid equations:

$$\sigma(\mathbf{x},l)\,\triangle p(\mathbf{x},l) = -\left\langle \frac{\partial}{\partial x_i}\sigma'(\mathbf{x})\,\frac{\partial}{\partial x_i}p'(\mathbf{x})\right\rangle \tag{8}$$

$$m(\mathbf{x},l)\frac{d\mathbf{x}}{dt} = -\sigma(\mathbf{x},l)\nabla p(\mathbf{x},l) - \langle \sigma'(\mathbf{x})\nabla p'(\mathbf{x})\rangle - \left\langle m'\frac{d\mathbf{x}'(t)}{dt}\right\rangle \tag{9}$$

The subgrid terms $\left\langle \frac{\partial}{\partial x_i}\sigma'(\mathbf{x})\cdot\frac{\partial}{\partial x_i}p'(\mathbf{x})\right\rangle$, $\left\langle m'\frac{d\mathbf{x}'(t)}{dt}\right\rangle$ in equations (8), (9) are unknown. This terms cannot be neglected without preliminary estimation since the correlation between the field $\sigma'(\mathbf{x})$, $m'(\mathbf{x})$ and the derivatives $p'(\mathbf{x})$, $\mathbf{x}'(t)$ may be significant. Subtracting equations (8), (9) from equations (1), (2) and taking into account only the first order terms, we obtain the subgrid equations:

$$\Delta p'(\mathbf{x}) = -\frac{1}{\sigma(\mathbf{x},l)}\frac{\partial}{\partial x_i}\left(\sigma'(\mathbf{x})\,\frac{\partial}{\partial x_i}p(\mathbf{x},l)\right), \tag{10}$$

$$m(\mathbf{x},l)\frac{d\mathbf{x}'(t)}{dt} = -\sigma(\mathbf{x},l)\nabla p'(\mathbf{x}) - \sigma'(\mathbf{x})\nabla p(\mathbf{x},l) - m'(\mathbf{x})\frac{d\mathbf{x}(t,l)}{dt}. \tag{11}$$

To evaluate subgrid terms in equations (8), (9), the right-hand sides of equations (10),(11) are considered to be known. This approach are described in detail in [4], [2], [5]. We suppose $\sigma(\mathbf{x},l)$, $p(\mathbf{x},l)$, $x(t,l)$ and their derivatives varying slower than, $\sigma'(\mathbf{x})$, $p'(\mathbf{x})$, $x'(t$ and their derivatives (the property of multiplicative cascades). Using the Green function, we can approximate the solution of equation (10):

$$p'(\mathbf{x},l) \approx \frac{1}{4\pi\sigma(\mathbf{x},l)}\int\limits_V \frac{1}{r}\frac{\partial}{\partial x'_j}\sigma'(\mathbf{x}')\,d\mathbf{x}'\ \frac{\partial p(\mathbf{x},l)}{\partial x'_j},\ r = |\mathbf{x}-\mathbf{x}'|\,. \tag{12}$$

The correlation functions are small outside the domain with radius $L \ll L_0$ and the center at the point $\mathbf{x}$, where L_0 is the scale of the domain V. Thus, integration over the finite volume V is changed to integration with an infinite limit. Using (12), formula $\frac{\partial^2}{\partial x_i \partial x_j}\frac{1}{r} = -\frac{1}{3}\delta_{ij}$, integrating by parts and changing the Cartesiasian coordinates to spherical coordinates, we obtain, that

$$\left\langle \sigma'(\mathbf{x})\frac{\partial p'(\mathbf{x})}{\partial x_i}\right\rangle \approx -\frac{1}{3}\Phi_0^{\varphi\varphi}(l)\,\frac{dl}{l}\,\sigma(\mathbf{x},l)\,\frac{\partial p(\mathbf{x},l)}{\partial x_i}. \tag{13}$$

In the same manner we can evaluate the correlation between $m'(\mathbf{x})$ and $\frac{\partial p'(\mathbf{x})}{\partial x_i}$:

$$\langle m'(\mathbf{x})\frac{\partial p'(\mathbf{x})}{\partial x_i}\rangle \approx -\frac{1}{3}\Phi_0^{\varphi\chi}(l)\,\frac{dl}{l}m(\mathbf{x},l)\,\frac{\partial p(\mathbf{x},l)}{\partial x_i}. \tag{14}$$

For the second subgrid term in (9) we have from the equation (11):

$$\left\langle m'(\mathbf{x})\frac{\mathbf{x}'(t)}{dt}\right\rangle = -\frac{\langle m'(\mathbf{x})\sigma'(\mathbf{x})\rangle}{m(\mathbf{x},l)}\nabla p(\mathbf{x},l) - \frac{\sigma(\mathbf{x},l)}{m(\mathbf{x},l)}\langle m'(\mathbf{x})\nabla p'(\mathbf{x})\rangle$$
$$-\frac{\langle m'(\mathbf{x})m'(\mathbf{x})\rangle}{m(\mathbf{x},l)}\frac{d\mathbf{x}(t,l)}{dt}. \tag{15}$$

Substituting (13), (14), (15) in equations (8), (9), we get

$$\nabla\left\{\sigma(\mathbf{x},l)\left[1-\frac{\Phi_0^{\varphi\varphi}}{3}\frac{dl}{l}\right]\right\}\nabla p(\mathbf{x},l)\approx 0, \tag{16}$$
$$m(\mathbf{x},l)\left[1-\Phi_0^{\chi\chi}\frac{dl}{l}\right]\frac{d\mathbf{x}(t,l)}{dt}\approx\sigma(\mathbf{x},l)\left[\left(\frac{\Phi_0^{\varphi\varphi}}{3}+\frac{2}{3}\Phi_0^{\varphi\chi}\right)\frac{dl}{l}-1\right]\nabla p(\mathbf{x},l).$$

The effective coefficients $\sigma(\mathbf{x})_l^{ef}$ and $m(\mathbf{x})_l^{ef}$ are evaluated by the formulas:

$$\sigma(\mathbf{x})_l^{ef}=\sigma_{0l}\exp\left[-\int_l^L\varphi(\mathbf{x},l_1)\frac{dl_1}{l_1}\right], m(\mathbf{x})_l^{ef}=m_{0l}\exp\left[-\int_l^L\chi(\mathbf{x},l_1)\frac{dl_1}{l_1}\right],$$

$$\sigma_{0l}\approx\sigma_0\left[1+\left(-\langle\varphi\rangle+\frac{1}{2}\Phi_0^{\varphi\varphi}\right)\frac{dl}{l}\right]\left[1-\frac{\Phi_0^{\varphi\varphi}}{3}\frac{dl}{l}\right]\approx\sigma_0\left[1+\left(-\langle\varphi\rangle+\frac{1}{6}\Phi_0^{\varphi\varphi}\right)\frac{dl}{l}\right],$$
$$m_{0l}\approx m_0\left[1-\Phi_0^{\chi\chi}\frac{dl}{l}\right],$$
$$\sigma_{0l}^1=\sigma_0\left[1+\left(-\langle\varphi\rangle+\frac{1}{2}\Phi_0^{\varphi\varphi}\right)\frac{dl}{l}\right]\left[\left(\frac{\Phi_0^{\varphi\varphi}}{3}+\frac{2}{3}\Phi_0^{\varphi\chi}\right)\frac{dl}{l}-1\right]\approx$$
$$\approx-\sigma_0\left[1+\left(-\langle\varphi\rangle+\frac{1}{6}\Phi_0^{\varphi\varphi}-\frac{2}{3}\Phi_0^{\varphi\chi}\right)\frac{dl}{l}\right].$$

For $dl\to 0$ we obtain

$$\frac{d\ln\sigma_{0l}}{d\ln l}=-\langle\varphi\rangle+\frac{1}{6}\Phi_0^{\varphi\varphi},\ \sigma_{0l_0}=\sigma_0,\ \frac{d\ln m_{0l}}{d\ln l}=-\Phi_0^{\chi\chi},\ \ m_{0l_0}=m_0,$$
$$\frac{d\ln\sigma_{0l}^1}{d\ln l}=-\langle\varphi\rangle+\frac{1}{6}\Phi_0^{\varphi\varphi}-\frac{2}{3}\Phi_0^{\varphi\chi},\ \sigma_{0l_0}^1=\sigma_0. \tag{17}$$

In scale invariant media the solution of the equations (17) have the form:

$$\sigma_{0l}=\sigma_0\left(\frac{l}{l_0}\right)^{\langle-\varphi\rangle+\frac{1}{6}\Phi_0^{\varphi\varphi}},\ m_{0l}=m_0\left(\frac{l}{l_0}\right)^{-\Phi_0^{\chi\chi}},\varepsilon_{0l}^1=\varepsilon_0\left(\frac{l}{l_0}\right)^{-\langle\varphi\rangle+\frac{1}{6}\Phi_0^{\varphi\varphi}-\frac{2}{3}\Phi_0^{\varphi\chi}}. \tag{18}$$

The variance of velocity vector in the scale-invariant medium is evaluated by the formulas:

$$\left\langle\left(\frac{d\mathbf{x}(t)}{dt}\right)^2\right\rangle-\left\langle\frac{d\mathbf{x}(t)}{dt}\right\rangle^2=\Omega\left[\frac{1}{3}\mathcal{A}-\Omega_1 E\right]\left(\frac{d\mathbf{x}(t,l)}{dt}\right)^2, \tag{19}$$

where $\Omega=\left(\frac{l}{l_0}\right)^{-\frac{4\Phi_0^{\varphi\chi}}{3}+\frac{2\Phi_0^{\varphi\varphi}}{3}+\Phi_0^{\chi\chi}}$ and $\Omega_1=\left(\frac{l}{l_0}\right)^{-2\langle\varphi\rangle-\frac{\Phi_0^{\varphi\varphi}}{3}+\Phi_0^{\chi\chi}}$, the main diagonal elements of matrix $\mathcal{A}$ equal $2\left(\frac{l}{l_0}\right)^{-\left(\Phi_0^{\varphi\varphi}/5\right)}+1$ and all other array elements are equal to $-\left(\frac{l}{l_0}\right)^{\left(-\Phi_0^{\varphi\varphi}/5\right)}+1$.

3 Numerical modeling

The equations (1), (2) are numerically solved in a cube L_0 on edge. On the sides of the cube $x_2 = 0$, $x_2 = L_0$ the pressure is constant $p(x_1, x_2, x_3)|_{x_2=0} = p_1$, $p(x_1, x_2, x_3)|_{x_2=L_0} = p_2$, $p_1 > p_2$. On the other sides of the cube, the pressure is specified by the linear relation for x_2: $p = p_1 + (p_2 - p_1)y/L_0$. The basic filtration flow is directed along x_2:-axis. For the numerical calculation the dimensionless variables are used. The all lengths are measured in terms of L_0, the unit difference of pressure is chose difference $p_1 - p_2$, the conductivity is measured in terms of σ_0. Thus, the problem is solved in unit cube with unit differential pressure for $\sigma_0 = 1$. For the spatial variables, we use 512^3 grid. At the first, the conductivity and the porosity fields are modelled. The integrals in (3), (5) are replaced by a finite difference formula, in which it is convenient to pass to the logarithm with base 2:

$$\sigma(\mathbf{x})_{l_0} = \exp\left[-\ln 2 \sum_{k=-8}^{0} \varphi(\mathbf{x}, \tau_k)\delta\tau\right], m(\mathbf{x})_{l_0} = m_0 \exp\left[-\ln 2 \sum_{k=-8}^{0} \chi(\mathbf{x}, \tau_k)\delta\tau\right]. \tag{20}$$

$$\varphi(\mathbf{x}, \tau_k) = \sqrt{\frac{\Phi_0^{\varphi\varphi}}{\ln 2}}\zeta(\mathbf{x}, \tau_k) + \langle\varphi\rangle, \chi(\mathbf{x}, \tau_k) = \sqrt{\frac{\Phi_0^{\chi\chi}}{\ln 2}}\left(r\zeta(\mathbf{x},\tau_k) + \sqrt{1-r^2}\varsigma(\mathbf{x},\tau_k)\right) + \frac{\Phi_0^{\chi\chi}}{2}. \tag{21}$$

Here is, $l_k = 2^{\tau_k}$, $\tau_k = k\delta\tau$, $\delta\tau = 1$ is discretization step in the logarithm of the scale, $-1 \le r \le 1$ is a correlation coefficient. The fields $\varphi(\mathbf{x}, \tau_k)$, $\chi(\mathbf{x}, \tau_k)$ are generated separately for each l. The total power exponents in (20) are summed over statistically independent layers. The number of additives in (20) and scales were chosen so that the largest scale of displacement pulsations would allow us to replace the probabilistic mean values by spatially average values, while the smallest scale of pulsations would provide that the numerical methods would approximate equations (1), (2) quite well on all the scales. Thus, three layers are used in the calculations $l_j = 8h, 16h, 32h$. A minimal scale is $l_0 = 1/64$. The independent Gaussian fields $\zeta(\mathbf{x}, \tau_k)$, $\varsigma(\mathbf{x},\tau_k)$ have unit variance, zero mean and correlation function $e^{\left[-(\mathbf{x}-\mathbf{y})^2\delta_{ij}/2^{2\tau_i}\right]}$. For the modeling the random fields the method [6] is used . For solving equation (1) an iterative method combined with the Fourier transform and the sweep method [7] and for solving equations (2) the Runge-Kutta method of second order accuracy are used. We use parameters:$m_0 = 0.15$, $\Phi_0^{\chi\chi} = 0.1$, $\langle\varphi\rangle = 0.2$, $\Phi_0^{\varphi\varphi} = 0.4$, $\langle\chi\rangle = 0.05$, $r = 0.8$, $\Phi_0^{\varphi\chi} = \sqrt{\Phi_0^{\varphi\varphi}\Phi_0^{\chi\chi}}$. For evaluation the mean velocity of the front is applied formula

$$\log_2\left\langle\frac{dx_2(\tau_k)}{dt}\right\rangle = a_k,\ a_k = \left(\langle\varphi\rangle - \frac{1}{6}\Phi_0^{\varphi\varphi} + \frac{2}{3}\Phi_0^{\varphi\chi} - \Phi_0^{\chi\chi}\right)k, \tag{22}$$

where $k = 1, 2, 3$ is the number of layers in (20). In the figure 1 The dashed line is the result obtained by the conventional perturbation theory, solid line is the result obtained by the effective equations, stars is the result of numerical modeling. In figure 2 the variance of $x_2(t)$ in double logarithmic coordinates is described. The increase in square thickness occurs according to a power law with an exponent 1.77. The exponent depending on time of particle size domain is equal to $0.87 > 0.5$, ($\sqrt{D_{x_2}(t)} \propto t^{0.87}$), that indicates the super diffusion process. In figure 3 the theoretical results described by (22) are compared with result of direct 3D numerical modeling and the result obtained by conventional perturbation theory.

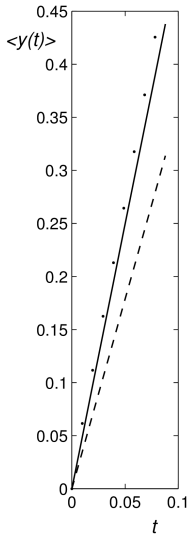


Figure 1: Mean velocity $\langle x_2(t)\rangle$.

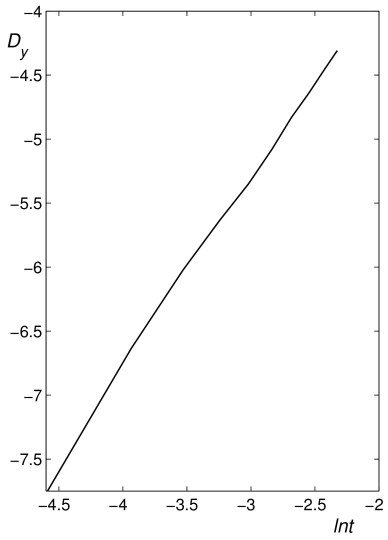


Figure 2: The dependence $D_{x_2} = \ln \langle (x_2(t) - \langle x_2(t)\rangle)^2\rangle$ on $\ln t$.

Conclusions

We have presented the effective coefficients for the wave equation if its parameters are described by extremely irregular small-scale fields that are close to multifractals. The multifractals can be obtained if a minimum scale l_0 in formulas (3), (5), tends to zero. As a minimum scale is finite, any singularities are absent, we use only the theory of differential equations and the theory of stochastic processes. The numerical verify, when use the spatial averaging and following additional averaging over Gibbs ensemble, gives good agreement with the theoretical results.

Acknowledgements

The Siberian Branch of the Russian Academy of Sciences (SB RAS) The work was supported by Program of Fundamental Scientific Research of State Academies of SciencesN0315-2015-0002, Russia.

References

[1] Monin A.S. and Yaglom A.M. (1975). *Statistical Fluid Mechanics*. Vol.**1**, MIT Press, Cambridge.

[2] Kuz'min G.A. , Soboleva O.N. (2002). Subgrid Modeling of Filtration in Porous Self-Similar Media. *Appl. Mech. Tech. Phys*. Vol.**43** , pp. 583-592.

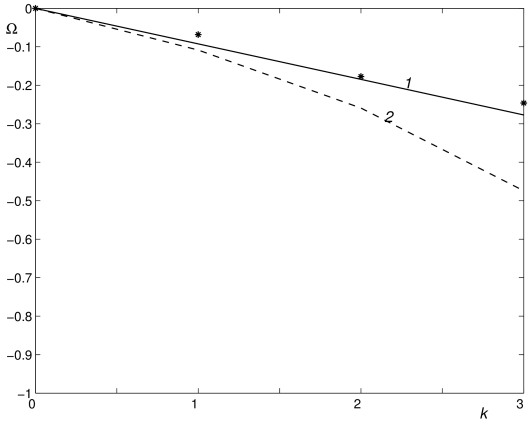


Figure 3: The logarithm of mean velocity $\Omega = \log_2 \left\langle \frac{dx_2}{dt} \right\rangle$.

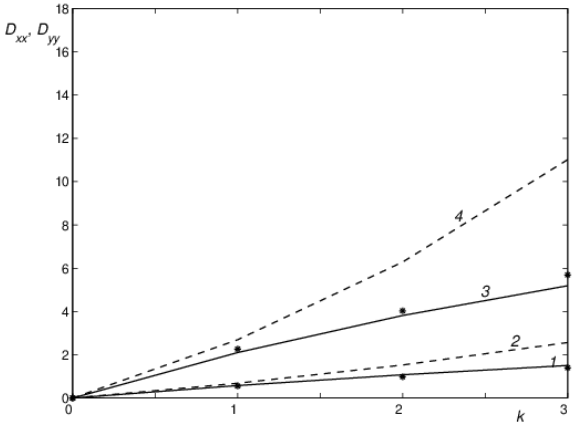


Figure 4: The variance of velocity vector $D_{x_1x_1}$ is the variance of x_1 axis velocity, lines 1,2; $D_{x_2x_2}$ is the variance of x_2 axis velocity, lines 3,4

[3] Kolmogorov A.N. (1962). A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *J. Fluid Mech.*, Vol. **13**, pp. 82–85.

[4] Dagan G. (1993). Higher order correction of effective permeability of heterogeneous isotropic formations of lognormal conductivity distribution. *Transport in Porous Media*. Vol.**12**,pp. 279–290.

[5] Shvidler M.I. (1985). *Statistical Hydrodynamics of Porous Media*. Nedra, Moscow, [in Russian].

[6] Ogorodnikov V.A., Prigarin S. M. (1996). *Numerical Modeling of Random Processes and Fields: Algorithms and Applications*. Utrecht, The Netherlands.

[7] Marchuk G.I.(1989) *Methods of Computational Mathematics*. Nauka, Moscow, [in Russian].

Maximum cross section method in estimation of jump-diffusion random processes

Tatyana Averina[1,2] and Konstantin Rybakov[3]
[1] *ICM&MG SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Moscow Aviation Institute (National Research University), Moscow, Russia*
e-mail: tatyana.averina@sscc.ru, rkoffice@mail.ru

Abstract

The paper deals with the estimation problem for random processes in dynamical systems whose mathematical models are described by stochastic differential equations with a compound Poisson process. The particle method and the maximum cross section method are applied for estimating (filtering, smoothing, and prediction).

Keywords: stochastic differential equation, Poisson process, estimation, filtering, smoothing, prediction, particle method, maximum cross section method, statistical modeling.

Introduction

In many practical problems [6, 8, 13, 16, 19, 20], such as the tracking, navigation, and financial analysis, it is necessary to estimate parameters or the state vector of a dynamical system given the observations. In this paper, the estimation problem is considered for jump-diffusion systems [9, 18].

The estimation of parameters or the state vector of the jump-diffusion system is considered for the current time (filtering problem) and for the past or future (smoothing or prediction, respectively) [7, 17]. The particle method, which involves modeling trajectories of a dynamical system, is used for solving these problems. Previously, the particle method was already applied to solve filtering problem for jump-diffusion systems [15], but the simplest method was used to simulate points of the Poisson process (the mathematical model of the jump-diffusion system involves the stochastic differential equation (SDE) with a compound Poisson process). Here, it is suggested to use the maximum cross section method and its modification [11, 12] for modeling these points in estimation algorithms. These methods provide more accurate estimation in the filtering, smoothing, and prediction problems. The complexity of these methods and their constructive dimension is lower due to fewer random number generator calls.

1 Estimation problem for jump-diffusion systems

We consider a stochastic continuous-time observation system that includes two equations:

$$dX(t) = f\big(t, X(t)\big)dt + \sigma\big(t, X(t)\big)dW(t) + \int_{\Xi} v\big(t, X(t-), \xi\big)\nu(dt \times d\xi), \qquad X(0) = X_0, \quad (1)$$

$$dY(t) = c\big(t, X(t)\big)dt + \zeta(t)dV(t), \quad Y(0) = Y_0 = 0. \quad (2)$$

The equation (1) is the Itô SDE with a compound Poisson process, where $X \in \mathbb{R}^n$ is a state vector, $t \in \mathbb{T} = [0, T]$, $f(t, x)\colon \mathbb{T} \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma(t, x)\colon \mathbb{T} \times \mathbb{R}^n \to \mathbb{R}^{n\times s}$, $v(t, x, \xi)\colon \mathbb{T} \times \mathbb{R}^n \times \Xi \to \mathbb{R}^n$, $\Xi = \mathbb{R}^q$; $W(t)$ is the s-dimensional standard Wiener process, ν is the Poisson random measure on $T \times \Xi$ with the characteristic measure Π_ν given by the function $\pi(t, x, \xi)\colon \mathbb{T} \times \mathbb{R}^n \times \Xi \to \mathbb{R}_+$; X_0 is an initial state vector with a given distribution.

The equation (2) is also SDE, where $Y \in \mathbb{R}^m$ is an observation, $c(t, x)\colon \mathbb{T} \times \mathbb{R}^n \to \mathbb{R}^m$, $\zeta(t)\colon \mathbb{T} \to \mathbb{R}^{m\times d}$; $V(t)$ is the d-dimensional standard Wiener process. The matrix $\eta(t) = \zeta(t)\zeta^{\mathrm{T}}(t)$ is nondegenerate, i.e., $\det \eta(t) \neq 0$ for any $t \in \mathbb{T}$. The initial state vector X_0, Wiener processes $W(t)$ and $V(t)$, the Poisson random measure ν are independent.

Let $\lambda(t, x)\colon \mathbb{T} \times \mathbb{R}^n \to \mathbb{R}_+$ denote the jump rate (or intensity) and let $\psi(t, \delta)\colon \mathbb{T} \times \mathbb{R}^n \to \mathbb{R}_+$ denote the probability density function for jumps (random increments of the state vector). These two functions define the characteristic measure Π_ν, the Poisson random measure ν, and the function $v(t, x, \xi)$. Thus,

$$\Pr\big(P(t + \Delta t) - P(t) = 1 \,|\, X(t) = x\big) = \lambda(t, x)\Delta t + o(\Delta t)$$

for small Δt, where Pr is a probability, $P(t)$ is the Poisson process such that [9, 10]

$$P(t) = \int_0^t \int_{\Xi} \nu(dt \times d\xi) = \nu([0, t] \times \Xi),$$

and $\psi(\tau_j, \delta)$ is the probability density function for the jump Δ_j at time τ_j, $\{\tau_j\}$ are points of the Poisson process $P(t)$, $\tau_0 = 0$, i.e.,

$$X(\tau_j) = X(\tau_j^-) + \Delta_j, \quad j = 1, 2, \ldots, \quad (3)$$

$$P^c(t) = \int_0^t \int_{\Xi} v\big(t, X(t^-), \xi\big)\nu(dt \times d\xi) = \sum_{j=1}^{P(t)} \Delta_j,$$

where $P^c(t)$ is the compound Poisson process.

Functions $f(t, x)$, $\sigma(t, x)$, $\lambda(t, x)$, $\psi(t, \delta)$, $c(t, x)$, and $\zeta(t)$ are given, they satisfy the conditions on the existence and uniqueness of the solution of SDEs [18]. In addition, $\mathrm{E}|X_0|^2 < \infty$, where E is an expectation or mean.

The optimal estimation problem is to find an estimate $\hat{X}(\theta)$ given the observations $Y_0^t = \{Y(\tau),\ \tau \in [0,t]\}$ such that $\hat{X}(\theta) = \psi(\theta, Y_0^t)$, where the function $\psi(\theta,\,\cdot\,)$ satisfies for all $\theta \in \mathbb{T}$ the following condition:

$$\mathrm{E}\big[\big(X(\theta) - \hat{X}(\theta)\big)^{\mathrm{T}}\big(X(\theta) - \hat{X}(\theta)\big)\big] \to \min_{\psi(\theta,\,\cdot\,)}.$$

This implies that $\hat{X}(\theta) = \psi(\theta, Y_0^t) = \mathrm{E}\big[X(\theta)|Y_0^t\big]$.

For $\theta = t$ we have the filtering problem, for $\theta < t$ and $\theta > t$ we have the smoothing problem and the prediction problem, respectively.

2 Particle method for estimation of jump-diffusion random processes

The algorithm based on modeling a special random process with terminating and branching paths was proposed in [14] to solve the optimal filtering problem. Paths of such process are completely determined by the SDE (1), and the observations described by the SDE (2) affect on the terminating and branching rates (or intensities). Then, this algorithm was modified so that it was possible to solve not only the filtering problem but also the prediction problem [4]. The smoothing problem is more difficult in this way, therefore, we will use the particle method [5].

The weight function $\omega(t)$ should be defined in the particle method [5]. This is a random process, which depends on the process $X(t)$ determined by the SDE (1) and observations Y_0^t. For example, the estimate $\hat{X}(t)$ in the filtering problem for the diffusion random process is the normalized weighted mean, i.e.,

$$\hat{X}(t) = \frac{\mathrm{E}\big[\omega(t)X(t)\big]}{\mathrm{E}\omega(t)}.$$

In this paper it is proposed to apply the particle method in the filtering, smoothing, and prediction problems for the jump-diffusion random process. We will focus only on modeling paths for the jump-diffusion random process using the maximum cross section method, modeling paths of the weight function, and finding the optimal estimate.

Let $\{t_k\}$ be a discretization of the time interval $\mathbb{T}$ with a variable step size $h_k > 0$:

$$t_{k+1} = t_k + h_k, \quad k = 1, 2, \ldots, N; \quad t_0 = 0, \quad t_N = T.$$

Denote by X_k a discrete-time approximation of the random process $X(t)$ determined by a numerical method for the Itô SDE (1) without the compound Poisson process [3, 9, 13], i.e., the random vector X_k is an approximation of $X(t)$ at time t_k, and $\lambda(t,x) \equiv 0$. For example,

$$X_{k+1} = X_k + h_k f(t_k, X_k) + \sqrt{h_k}\,\sigma(t_k, X_k)\,\zeta_k \quad \text{(Euler–Maruyama method)}; \qquad (4)$$

$$X_{k+1} = X_k + \frac{h_k}{2}\big(a(t_k, X_k) + a(t_{k+1}, X_k^p)\big) + \frac{\sqrt{h_k}}{2}\big(\sigma(t_k, X_k) + \sigma(t_{k+1}, X_k^p)\big)\zeta_k \quad \text{(Derivative-free Heun method)}, \quad (5)$$

where $X_k^p = X_k + h_k\, f(t_k, X_k) + \sqrt{h_k}\,\sigma(t_k, X_k)\,\zeta_k$;

$$X_{k+1} = X_k + \frac{h_k}{2}\left[I - \frac{h_k}{2}\frac{\partial a(t_k, X_k)}{\partial x}\right]^{-1}\big(a(t_k, X_k) + a(t_k, X_k^p)\big) + \frac{\sqrt{h_k}}{2}\big(\sigma(t_k, X_k) + \sigma(t_k, X_k^p)\big)\zeta_k \quad \text{(Rosenbrock type method)}, \quad (6)$$

where I is the n-dimensional identity matrix, $X_k^p = X_k + \sqrt{h_k}\,\sigma(t_k, X_k)\,\zeta_k$, and

$$a_i(t,x) = f_i(t,x) - \frac{1}{2}\sum_{j=1}^{n}\sum_{l=1}^{s}\frac{\partial \sigma_{il}(t,x)}{\partial x_j}\,\sigma_{jl}(t,x), \quad i = 1, 2, \ldots, n.$$

In relations given above ζ_k is an s-dimensional random vector with independent coordinates having the standard normal distribution for all k.

According to [5] the weight function $\omega(t)$ is defined by

$$\omega(t) = \exp\left\{\int_0^t c^{\mathrm{T}}\big(\tau, X(\tau)\big)\eta^{-1}(\tau)dY(\tau) - \frac{1}{2}\int_0^t c\big(\tau, X(\tau)\big)\eta^{-1}(\tau)c\big(\tau, X(\tau)\big)d\tau\right\},$$

and a discrete-time approximation of the random process $\omega(t)$ at time t_k is denoted by ω_k, where

$$\omega_{k+1} = \omega_k \exp\left\{c^{\mathrm{T}}(t_k, X_k)\eta^{-1}(t_k)\big(Y(t_{k+1}) - Y(t_k)\big) - \frac{1}{2}\,c^{\mathrm{T}}(t_k, x_k)\eta^{-1}(t_k)c(t_k, X_k)h_k\right\}, \quad \omega_0 = 1. \quad (7)$$

The discretization $\{t_k\}$ is a superposition of the discretization of the time interval $\mathbb{T}$ with a fixed step size $h > 0$ and points of the Poisson process $P(t)$. It is suggested to use the maximum cross section method for modeling these points $\{\tau_j\}$. According to [1, 2, 11, 12], if there exists λ^* such that $\lambda(t) \leqslant \lambda^*$, then the random time between neighboring points τ_j and τ_{j+1} should be simulated as follows:

$$\tau = \theta_{\mathcal{N}}, \quad \mathcal{N} = \min\left\{\vartheta\colon\ \alpha_\vartheta \leqslant \frac{\lambda(\tau_j + \theta_\vartheta)}{\lambda^*}\right\}, \quad \theta_\vartheta = \sum_{i=1}^{\vartheta}\xi^i,$$

where $\xi^1, \xi^2, \ldots, \xi^\vartheta, \ldots$ is a sequence of independent random variables having the exponential distribution with parameter λ^*: $\xi^i = -\ln\beta_i/\lambda^*$; $\alpha_1, \alpha_2, \ldots, \alpha_\vartheta, \ldots$, $\beta_1, \beta_2, \ldots, \beta_\vartheta, \ldots$ is a sequence of independent random variables having the uniform distribution on the interval $(0, 1)$, and $\lambda(t) = \lambda(t, X(t))$.

The modified maximum cross section method [1, 2, 12] is more efficient, and for this modified method the number $\mathcal{N}$ is defined by

$$\mathcal{N} = \min\left\{\vartheta\colon\ 1-\alpha > \prod_{i=1}^{\vartheta}\left(1-\frac{\lambda(\tau_j+\theta_i)}{\lambda^*}\right)\right\},$$

where α is a random variable having the uniform distribution on the interval $(0,1)$.

To find the approximate solution of the filtering, smoothing, and prediction problems it is necessary to simulate M sample paths $X^i(t)$ of the random process $X(t)$ and the corresponding paths $\omega^i(t)$ of the weight function $\omega(t)$ by the scheme (4), (5) or (6) and the relation (7) taking into consideration points of the Poisson process $P(t)$ and the relation (3), $i = 1, 2, \ldots, M$.

The approximate solution of the optimal estimation problem is the normalized weighted mean:

$$\hat{X}(t_\kappa) \approx \hat{X}_\kappa = \left(\sum_{i=1}^{M}\omega_k^i\right)^{-1}\sum_{i=1}^{M}\omega_k^i X_\kappa^i,$$

where the index k corresponds to the current time $t = t_k$ and the index κ corresponds to the time $\theta = t_\kappa$ for which the state vector estimate is calculated. The higher order moments can be also found as well as estimations of the probability density function or distribution function of the state vector.

Thus, it is obtained the approximate solution of the filtering problem if $\kappa = k$. We have the approximate solution of the smoothing problem and the prediction problem if $\kappa < k$ and $\kappa > k$, respectively.

Note that in the smoothing problem unlike the filtering problem it is necessary to store paths of the random process $X(t)$ in the computer memory. In the filtering problem it is enough to store the states for the current time only, these states define the initial data in the prediction problem [4].

Acknowledgements

This work is financially supported in part by the base project ICM&MG SB RAS no. 0315-2016-0002 and the Russian Foundation for Basic Research, project no. 17-08-00530.

References

[1] Averina T.A. (2013). A modified algorithm for statistical simulation of random structure systems with distributed transitions. *Numer. Anal. Appl.* Vol. **6**, No. 2, pp. 91–97.

[2] Averina T.A. (2016). A randomized maximum cross-section method to simulate random structure systems with distributed transitions. *Numer. Anal. Appl.* Vol. **9**, No. 3, pp. 179–190.

[3] Averina T.A., Karachanskaya E.V., Rybakov K.A. (2018). Statistical analysis of diffusion systems with invariants. *Rus. J. Numer. Anal. Math. Modelling*. Vol. **33**, No. 1, pp. 1–13.

[4] Averina T.A., Rybakov K.A. (2017). An approximate solution of the prediction problem for stochastic jump-diffusion systems. *Numer. Anal. Appl.* Vol. **10**, No. 1, pp. 1–10.

[5] Bain A., Crisan D. (2009). *Fundamentals of Stochastic Filtering*. Springer, New York.

[6] Bar-Shalom Y., Li X.R., Kirubarajan T. (2001). *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons.

[7] Einicke G.A. (2012). *Smoothing, Filtering and Prediction: Estimating the Past, Present and Future*. InTech, Rijeka.

[8] Grewal M.S., Weill L.R., Andrews A.P. (2007). Global Positioning Systems, Inertial Navigation, and Integration. John Wiley & Sons.

[9] Hanson F.B. (2007). *Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis, and Computation*. SIAM, Philadelphia.

[10] Last G., Penrose M. (2017). *Lectures on the Poisson Process*. Cambridge University Press, Cambridge.

[11] Mikhailov G.A. (1991). *Minimization of Computational Costs of Non-Analogue Monte Carlo Methods*. World Scientific, Singapore.

[12] Mikhailov G.A., Averina T.A. (2009). The maximal section algorithm in the Monte Carlo method. *Dokl. Math.* Vol. **80**, No. 2, pp. 671–673.

[13] Platen E., Bruti-Liberati N. (2010). *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*. Springer, Berlin.

[14] Rybakov K.A. (2016). Statistical algorithms of optimal filtering problem for nonlinear jump-diffusion models. *UGATU Bulletin*. Vol. **20**, No. 4, pp. 107–113. (In Russian)

[15] Rybakov K.A. (2018). Continuous particle filters for stochastic jump-diffusion systems. *In Proc. of Int. Conf. on Actual Problems of Applied Mathematics, Computer Science and Mechanics*, 17–19 December, 2018, Voronezh, Russia, pp. 1613–1619. (In Russian)

[16] Rybakov K.A. (2019). Solving the nonlinear problems of estimation for navigation data processing using continuous particle filter. *Gyroscopy and Navigation*. Vol. **10**, No. 1, pp. 27–34.

[17] Särkkä S. (2013). *Bayesian Filtering and Smoothing*. Cambridge University Press, Cambridge.

[18] Situ R. (2005). *Theory of Stochastic Differential Equations with Jumps and Applications*. Springer, New York.

[19] Stepanov O.A., Toropov A.B. (2015). Nonlinear filtering for map-aided navigation. Part 1. An overview of algorithms. *Gyroscopy and Navigation*. Vol. **6**, No. 4, pp. 324–337.

[20] Stepanov O.A., Toropov A.B. (2016). Nonlinear filtering for map-aided navigation. Part 2. Trends in the algorithm development. *Gyroscopy and Navigation*. Vol. **7**, No. 1, pp. 82–89.

A stochastic model of an unmanned aerial vehicle control system

Tatyana A. Averina [1,2], Ivan M. Kosachev [3] and Konstantin N. Chugai [4]
[1] *Institute of Computational Mathematics and Mathematical Geophysics SB RAS*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Military Academy of the Republic of Belarus, Minsk, Republic of Belarus*
[4] *Research Institute of the Armed Forces, Minsk, Republic of Belarus*
e-mail: tatyana.averina@sscc.ru, kosachev1301@mail.ru, konstantin.ch40@gmail.com

Abstract

In this paper, a model of a random-structure system is constructed. The model describes control of an unmanned aerial vehicle (AV). The model is tested by using a specially designed statistical algorithm.

Keywords: unmanned aerial vehicle, block diagram, random structure system, stochastic differential equation, statistical algorithm

Introduction

At present, dynamic systems with random changes in the conditions of functioning and perturbations causing abrupt changes in the structure as a whole (that is, structural uncertainty) are widely used. A systematic description of such problems and methods of their analysis can be found in [3].

Such models have been developed in many branches of science to perform scientific research related to modeling of complicated phenomena and control processes. These are, for instance, problems of automatic control of a system with different operation modes and different structures in non-overlapping time intervals. Such systems are called dynamic systems with random structure change (DSRSC) or random-structure systems.

Examples of random-structure systems are airborne collision avoidance systems, systems of search for and interception of signals in navigation and AV flight control, combined target guidance systems, as well as systems with possible failures.

1 Construction of a stochastic model for an unmanned aerial vehicle control system

We construct a mathematical model of a control system (CS) of an unmanned aerial vehicle (AV) of the DSRSC class.

The system being simulated has two states: The first (basic) state is searching for a target and tracking by AV angular coordinates and distance using a radar information system. The second one is target interception and autotracking by the angular coordinates and distance using a laser locator. If the laser information system

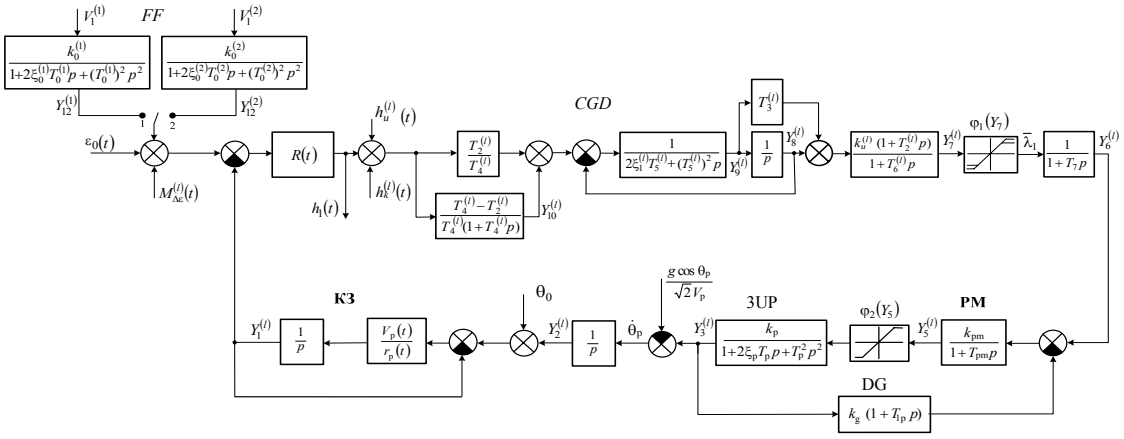


Figure 1: Block diagram of the guidance loop of a remote controlled AV in a vertical plane

fails, it rapidly changes to the radar information system, that is, the system returns to the first state.

Transitions from the r-th structure to the l-th one are governed by using transition intensities $\nu^{(r,l)}(t)$ or conditional probabilities $\mathrm{P}^{(r,l)}(t, A)$ calculated as convolutions of the probability density distribution (PDD) of the "life" times preceding the l-th structure.

The PDD of the DSRSC "life" times in each of the possible structures, $f^{(1)}(\tau)$ and $f^{(2)}(\tau)$, is determined from results of natural (or seminatural) experiments.

Fig. 1 presents a typical block diagram of the guidance loop of a distance controlled AV in a vertical plane with an integrated (radar-laser) information system. The following notation is used in this figure:

FF, transfer functions of the forming filters in the form of oscillatory links, providing the formation of the output error of measurement of the difference between the angular coordinates of the target and AV in a vertical plane in the l-th structure (phase coordinate);

R(t), deterministic function approximating the current slant distance to AV;

CGD, transfer function of the command guidance system of AV in a vertical plane;

PM, transfer function of the AV control actuator in the form of an aperiodic link;

3UP, transfer characteristic of the AV in the form of an oscillatory link with nonstationary parameters;

DG, transfer function of damping gyroscope;

K3, transfer function of nonstationary kinematic link of the AV remote controlled loop.

The following assumptions and constraints are used in the AV remote controlled loop block diagram shown in Fig. 1:

1) a two-dimensional model of the AV remote controlled loop is considered only in one (vertical) guidance plane;

2) strong radar and laser signals reflected from the target are considered, in which

the ratio of the signal power to the noise power is much larger than unity;

3) pulsed control commands transmitted to the AV are not taken into account, since the re-mote controlled loop is a narrow-band tracking system;

4) measurement errors of the difference in the angular coordinates of a single target and in vertical and horizontal planes are assumed to be Gaussian random processes with exponential-oscillatory correlation functions

$$K_y^{(l)}(\tau) = D_y^{(l)} e^{-\alpha^{(l)}|\tau|} cos(\beta^{(l)}|\tau|).$$

Table 1: Description of the phase coordinates of the AV control system

phase coord.	phase coordinate name	unit
$y_1^{(l)}(t)$	AV tilt angle at the kinematic link output in the l-th CS structure	rad
$y_2^{(l)}(t)$	Pitch angle in the l-th CS structure	rad
$y_3^{(l)}(t)$	Angular velocity of rotation of the AV velocity vector in a vertical plane in the l-th CS structure	$\frac{\text{rad}}{\text{s}}$
$y_4^{(l)}(t)$	Angular acceleration of rotation of the AV velocity vector in a vertical plane in the l-th CS structure	$\frac{\text{rad}}{\text{s}^2}$
$y_5^{(l)}(t)$	AV elevator deflection angle in the l-th CS structure	rad
$y_6^{(l)}(t)$	AV control command at the onboard command decoder output output in the l-th CS structure	rad
$y_7^{(l)}(t)$	AV control command in a vertical plane at the CGD output (up to the stopper) sent to the AV in the l-th CS structure	m
$y_8^{(l)}(t)$	Signal at the output of the second CGD forming filter in the l-th CS structure	m
$y_9^{(l)}(t)$	Signal derivative at the output of the second CGD forming filter in the l-th CS structure	$\frac{\text{m}}{\text{s}}$
$y_{10}^{(l)}(t)$	Signal at the output of the first CGD forming filter in the l-th CS structure	rad
$y_{11}^{(l)}(t)$	Error derivative of measurement of the difference between the angles of target location and AV at the FF output in the l-th CS structure	$\frac{\text{rad}}{\text{s}}$
$y_{12}^{(l)}(t)$	Error of measurement of the difference between the angles of target location and AV at the FF output in the l-th structure	rad

When investigating only the AV remote controlled loop this allows replacing the radar and laser systems of measuring the difference in the angular coordinates of the target and vehicle by two equivalent forming filters in the form of oscillatory links. The parameters of these forming filters are calculated using the values of the

parameters $D^{(l)}$, $\alpha^{(l)}$, $\beta^{(l)}$, in the exponential-oscillation correlation functions for every structure of the control system.

The input master control in the AV remote controlled loop is the current angle of target location in a vertical plane ε_0.

The deterministic function $M^{(l)}_{\Delta\varepsilon}(t)$ takes into account the nonstationary mean error in the measurement of the difference of the angular coordinates of the target and AV in a vertical guiding plane in the l-th structure.

The deterministic function $h^{(l)}_u(t)$ is used to form a preemptive kinematic trajectory of AV flight with respect to the target sighting line.

The function $h^{(l)}_k(t)$ compensates for the dynamic error of the AV guidance loop in the l-th structure.

The function $gcos(\theta_p)/(\sqrt{2}V_p)$ takes into account the effect of the AV mass on the pitch angle variation rate (the angle of inclination of the AV velocity vector to the horizon).

Phase coordinates are introduced only at the outputs of inertial links (integrators, aperiodic and oscillatory ones). The physical meaning and dimensions of the phase coordinates (see Fig. 1) of the AV control system under study is explained in the table.

Using standard rules, a random-structure system is constructed for each l-th ($l = 1, 2$) state of the structure:

$$\frac{dy^{(l)}_1}{dt} = \frac{V_p(t)}{r_p(t)}(\theta_0 - y^{(l)}_1 + y^{(l)}_2),$$

$$\frac{dy^{(l)}_2}{dt} = -g\frac{cos(\theta_p)}{\sqrt{2}V_p(t)} + y^{(l)}_3,$$

$$\frac{dy^{(l)}_3}{dt} = y^{(l)}_4,$$

$$\frac{dy^{(l)}_4}{dt} = -\frac{1}{T^2_p(t)}y^{(l)}_3 - \frac{2\xi_p(t)}{T_p(t)}y^{(l)}_4 + \frac{k_p(t)}{T^2_p(t)}\varphi^{(l)}_2(y_5),$$

$$\frac{dy^{(l)}_5}{dt} = -\frac{k_g k_{pm}}{T_{pm}}(y^{(l)}_3 + T_{1p}y^{(l)}_4) + \frac{1}{T_{pm}}(k_{pm}y^{(l)}_6 - y^{(l)}_5 + k_{pm}V_n(t)),$$

$$\frac{dy^{(l)}_6}{dt} = \frac{1}{T_7}(-y^{(l)}_6 + k^{(l)}_u\varphi^{(l)}_1(y_7) + V_n(t)) + \frac{k^{(l)}_u T^{(l)}_2}{T^{(l)}_6 T_7}\varphi^{(l)}_1(y_7)(y^{(l)}_8 + T^{(l)}_3 y^{(l)}_9),$$

$$\frac{dy^{(l)}_7}{dt} = \frac{k^{(l)}_u (T^{(l)}_2)^2 T^{(l)}_3}{T^{(l)}_4 (T^{(l)}_5)^2 T^{(l)}_6}\Big(R(t)[\varepsilon_0(t) + M^{(l)}_{\Delta\varepsilon} - y^{(l)}_1 + y^{(l)}_{12}] + h^{(l)}_u + h^{(l)}_k\Big) +$$
$$\frac{k^{(l)}_u (T^{(l)}_5)^2 - T^{(l)}_2 T^{(l)}_3 T^{(l)}_6}{(T^{(l)}_5)^2 T^{(l)}_6} y^{(l)}_8 - \frac{1}{T^{(l)}_6}y^{(l)}_7 + \frac{k_u(l)(T^{(l)}_2 + T^{(l)}_3)T^{(l)}_5 - 2\xi_1 T^{(l)}_2 T^{(l)}_3}{T^{(l)}_5 T^{(l)}_6} y^{(l)}_9 + \frac{k^{(l)}_u T^{(l)}_2 T^{(l)}_3}{(T^{(l)}_5)^2 T^{(l)}_6} y^{(l)}_{10},$$

$$\frac{dy^{(l)}_8}{dt} = y^{(l)}_9,$$

$$\frac{dy^{(l)}_9}{dt} = \frac{T^{(l)}_2}{T^{(l)}_4 (T^{(l)}_5)^2}\Big(R(t)[\varepsilon_0(t) + M^{(l)}_{\Delta\varepsilon} - y^{(l)}_1 + y^{(l)}_{12}] + h^{(l)}_u + h^{(l)}_k\Big) - \frac{2\xi^{(l)}_1}{T^{(l)}_5}y^{(l)}_9 + \frac{1}{(T^{(l)}_5)^2}(y^{(l)}_{10} - y^{(l)}_8),$$

$$\frac{dy_{10}^{(l)}}{dt} = \frac{T_4^{(l)} - T_2^{(l)}}{(T_4^{(l)})^2}[R(t)(\varepsilon_0(t) + M_{\Delta\varepsilon}^{(l)} - y_1^{(l)} + y_{12}^{(l)}) + h_u^{(l)} + h_{k1}^{(l)}] - \frac{1}{T_4^{(l)}} y_{10}^{(l)},$$

$$\frac{dy_{11}^{(l)}}{dt} = -\frac{1}{(T_0^{(l)})^2} y_{11}^{(l)} - \frac{2\xi_0^{(l)}}{T_0^{(l)}} y_{11}^{(l)} + \frac{k_0^{(l)}}{(T_0^{(l)})^2} V_1(t),$$

$$\frac{dy_{12}^{(l)}}{dt} = y_{11}^{(l)},$$

These equations are nonstationary and nonlinear with additive white noises V_n, V_1. Nonlinearity: $\varphi_1^{(l)}(y_7)$ limits the control command size, and $\varphi_2^{(l)}(y_5)$ limits the angle of deviation of the AV steering wheels.

This system can be written as a system with random structure given by stochastic differential equations:

$$dY(t) = F^{(l)}(t, Y(t))dt + \Sigma(t)^{(l)} dw(t),$$

where $w(t)$ is 2-dimensional standard Winer process.

The input master control for the AV control system under study is the current target location angle $\varepsilon_0(t)$, described by the expression

$$\varepsilon_0(t) = \arctan \frac{y_0(t)}{x_0(t)} = \arcsin \frac{y_0(t)}{r_0(t)},$$

where $y_0(t)$ is the current target flight altitude; $r_0(t) = \sqrt{x_0^2(t) + y_0^2(t)}$ is the current slant distance to the target in a vertical plane; and $x_0(t)$ is the current horizontal distance to the target.

2 Statistical modeling of the control system

When changing the structures, some natural conditions of reconstruction are considered. The exception is $y_{12}^{(l)}(t)$, which is reconstructed in the neighborhood of the functions $M_{\Delta\varepsilon}^{(1)}(t)$ and $M_{\Delta\varepsilon}^{(2)}(t)$).

LA error in the vertical plane in the l-th structure is calculated by the formula

$$h_1^{(l)} = (\varepsilon_0 + M_{\Delta\varepsilon}^{(l)} - y_1^{(l)} + y_{12}^{(l)})R(t_v),$$

where t_v is the point of contact. An indicator of the efficiency of the AV control system is the probability of getting into a circle of radius R_{PB} specified with respect to the target. Under a normal PDD of the error, $f(h_1^{(l)})$ is calculated by the following formula:

$$\mathrm{P}(-R_{PB} \le h_1 \le R_{PB}) = \int_{-PB}^{PB} f(h_1)dh_1 = \Phi\Big(\frac{R_{PB} - M_{h_1}}{\sigma_{h_1}}\Big) + \Phi\Big(\frac{R_{PB} + M_{h_1}}{\sigma_{h_1}}\Big),$$

where h_1 is the unconditional (taking into account the two states of the structure) AV error in a vertical plane; M_{h_1} is the unconditional mean of the AV error in a vertical guidance plane; σ_{h_1} is the unconditional standard deviation of the AV error in a vertical guidance plane; and $\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_0^u e^{-\frac{t^2}{2}} dt$ is the Laplace function.

Statistical modeling of the solutions of system (1) allows estimating various probabilistic characteristics of the solution, including the PDD error. The above-developed stochastic model of the DSRSC class was studied on an Intel Core i5 3330 PC (3.00 GHz) using an algorithm for statistical modeling of random-structure systems with distributed transitions described in [1] and stochastic Euler method

$$Y_{n+1} = Y_n + F^{(l)}(t_n, Y_n)h + \Sigma^{(l)}(t_n)\sqrt{h}\zeta_n,$$

where ζ_n are independent standard Gaussian variables, $n = 1, 2, ...,$

$$\zeta_{2s-1} = \sqrt{-2\ln\alpha_s}\cos(2\pi\beta_s),\ \zeta_{2s} = \sqrt{-2\ln\alpha_s}\sin(2\pi\beta_s),\ s = 1, 2, ...$$

The RND128 pseudorandom generator [5] (with a modulus 2^{128} and a multiplier 5^{100109}) was used for the simulation of uniform random variables α_s, β_s on the interval $(0, 1)$. $N = 10^5$ samples of the solution were used in numerical examples. The time mesh nodes include a uniform mesh with the step $h = 0.1$ and the moments of change of structure.

The constant parameters of the model are shown in table 2.

Table 2: The constant parameters of the model

	if $t \leq 21.2$	if $t > 21.2$
k_0	0.591	-0.039
k_1	-0.0147	0.016
k_2	0.537	-0.2364
k_3	-0.014	0.022
k_4	0.06857	-0.0261
k_5	-0.0018	0.0026
k_6	0.1205	-0.0063
k_7	-0.0015	0.0044
k_8	551	2662
k_9	14.1666	-108.53
k_{10}	-0.1334	0.9559
k_{11}	-834	-21494

The following nondimensional parameters of the model were specified for the test calculations:

$$k_0 = 0,707,\ \xi_0 = 0,27,\ T_0^{(1)} = 0.76,\ T_0^{(2)} = 0.37,\ k_u^{(1)} = 1,\ k_u^{(2)} = 3,\ \xi_1 = 0.8,$$
$$k_{pM} = 0.5,\ T_{pM} = 0.035,\ k_g = 0.375,\ T_2^{(1)} = 1.8,\ T_2^{(2)} = 1.0,\ T_3^{(1)} = 0.33,$$
$$T_3^{(2)} = 0.174,\ T_4^{(1)} = 0.14,\ T_4^{(2)} = 0.0725,\ T_5^{(1)} = 0.14,\ T_5^{(2)} = 0.063,$$
$$T_6^{(1)} = 5.0,\ T_6^{(2)} = 2.5,\ T_7 = 0,04,\ T_{1p}(t) = k_0 + k_1 t,$$
$$T_p(t) = k_4 + k_5 t,\ k_p(t) = k_2 + k_3 t,\ \xi_p(t) = k_6 + k_7 t,$$
$$r_p(t) = k_{11} + k_8 t + \tfrac{k_9}{2} t^2 + \tfrac{k_{10}}{3} t^3,\ V_p(t) = k_8 + k_9 t + k_{10} t^2.$$

Conclusions

In this paper, a stochastic model of a control system of unmanned aerial vehicles of the DSRSC class has been developed. Test calculations with this model using a statistical algorithm have shown that it can be successfully used with some nondimensional parameters of natural and semi-natural tests. It is planned to

- develop an analytical mathematical model with the same AV remote controlled loop according to the method [4];
- perform additional studies with a modified statistical algorithm [2] of the stochastic CS model constructed in the present paper;
- compare the results of calculations in analytical and statistical modeling.

Acknowledgements

This work is financially supported in part by the base project no. 0315-2016-0002.

References

[1] Averina T.A. (2013). A modified algorithm for statistical simulation of random structure systems with distributed transitions. *Numerical Analysis and Applications*. Vol. **6**, pp. 91-97.

[2] Averina T.A. (2016). A randomized maximum cross-section method to simulate random structure systems with distributed transitions. *Numerical Analysis and Applications*. Vol. **9**, pp. 179-190.

[3] Kazakov I.E., Artemiev V.M., Bukhalev V.A. (1993). *Analysis of Random Structure Systems*. Fizmatlit, Moscow.

[4] Kosachev I.M., Eroshenkov V.G. (1993). *Analytical modeling of stochastic systems*. Science and Technology, Minsk.

[5] Mikhailov G.A., Marchenko M.A. (2002). Parallel realization of statistical simulation and random number generators. *Russ. J. Numer. Anal. Math. Modelling*. Vol. **17**, pp. 113-124.

Evaluating the impact of tourism on economic growth in Tunisia

Shakra Moayyad[1] ,Shmidt Yuriy[1] and Almosabbeh Imadeeddin[2]
[1] *Far Eastern Federal University, Vladivostok, Russia*
[2] *Qassim University,Buraydah, Saudi Arabia*
e-mail: shakra.m@students.dvfu.ru, syd@dvfu.ru

Abstract

The main objective of this study is to examine the impact of tourism on the economic growth in Tunisia during the period (1985-2017). To achieve this goal, the ARDL model was used in its form reduced using two variables (tourism revenues (TOUR), exchange rate (EX)). The unit root test was also used to determine if the variables were stable over time and if the variable (LTOUR) was stable at the level, but the variables (LGDP, LEX) were stable after the first difference level. The results of the study showed a positive effect of the tourism revenues, but the effect of exchange rate was not significant in the long term.

Keywords: Tunisian economy, Tourism revenues, Autoregressive-Distributed Lag model (ARDL).

Introduction

According to the WTO, Holloway and others [3, 12], tourism activity is defined as the activities of people on their journey and residence in a place outside their habitual residence during a continuous period of less than one year for recreation, business or other purposes. Thus, income from tourism can be defined as follows: the total income from tourism activities during the year.

The focus of tourism research since the 1930s has been on the importance of tourism as a source of foreign exchange. But the contribution of tourism and its impact on development and macroeconomic variables have only recently been discussed. There is no doubt that tourism is the main engine of a country's economy and that it has a positive impact on the economy.

The World Travel and Tourism Council succinctly summarized the important role that travel and tourism play in the growth of global economy as follows:

In 2017, Travel & Tourism' s total contribution to the global economy has risen to 10.4% of global GDP (US $8.3 trillion), as it grows at a faster pace than most other important sectors such as trade, finance, transportation and manufacturing. In total, nearly 313 million jobs related to tourism were created, which means that 1 of 10 jobs around the world come from the travel and tourism sector. Travel and tourism still capable of generating high levels of employment through ever-increasing demand, demonstrating the importance and value of the sector as a tool for economic development and job creation [11].

In this paper, the impact of tourism on economic growth in Tunisia will be measured through the use of the Nicolas Detsakis's model.

1 Pilot study methodology

Cointegrating tests, such as Engle and Granger (1987), Johansen (1988) and Johansen and Juselius (1990), require that variables be in the same level: in this case they cannot be performed with integral variables of different levels, i.e. I(0) and I(1). Therefore, the Autoregressive-Distributed Lag (ARDL) model has become the best alternative to what does not require that the estimated variables have the same level of integration.
The ARDL method used to test the overall cointegrating has many advantages:
It can be applied regardless of whether the variables studied are cointegrated from I(0) or I(1) or cointegrated in different levels, that is, they can be used when the degree of cointegration is unknown or not homogeneous for all variables. This would be good if the sample size (number of observations) is small, and this is not the case for most traditional cointegration tests that require a large sample size so that the results are more effective. Moreover, its use allows us to simultaneously evaluate the components of the long-term and short-term (relations) in one equation instead of two separate equations [4].

The ARDL method will be used in three stages: at the first stage, cointegration test within the UECM, which adopts the following formula, imposing a relationship between Y (dependent variable) and X (independent variable vector):

$$\Delta Y_t = \alpha_0 + \sum_{i=1}^{m} \beta_0 \Delta Y_{t-i} + \sum_{i=0}^{n} \Theta_i \Delta X_{t-i} + \lambda_1 Y_{t-1} + \lambda_2 X_{t-1} + \eta_t \qquad (1)$$

Where λ_1, λ_2 expresses long-term ratios, and β, Θ expresses short-term ratios, and Δ denotes the first differences of variables, while each from (m, n) lag for variables (although not necessarily Number of time delay periods) for variables on the same level or quantity ($m \neq n$) [9]. And η the random error, which has an average equal to zero and a constant variance and does not have consecutive correlations between them.

Cointegration is checked between variables in equation (1) using the following assumptions: Null hypothesis (H0): no cointegration: $\lambda_1 = \lambda_2 = 0$, in comparison with the alternative hypothesis (H1): cointegration $\lambda_1 \neq \lambda_2 \neq 0$ Because the distribution of test F is non-standard and depends on: (1) whether the variables included in the ARDL form are cointegrated from I(0) or I(1); (2) the number of independent variables; (3) sample size, and therefore the null hypothesis is rejected by comparing the calculated F values with the values set within the critical limits proposed by Pesaran et al. (2001), Two sets of asymptotic critical values are provided: one when all regressors are I(1) and the other if they are all I(0). These two sets of critical values provide a band covering all possible classifications of the regressors into I(0), I(1) or mutually cointegrated [5]. If the calculated value for F is greater than UCB, in this case the null hypothesis is rejected and an alternative hypothesis (cointegration) is accepted. Conversely, if the calculated F is less than the LCB, in this case the null hypothesis is accepted (no cointegration). If the F value is between UCB and LCB, the result will not be adjusted.

In the case of general cointegration of variables, the second stage includes an assessment of the long-term equation as follows:

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \vartheta_i Y_{t-i} + \sum_{i=0}^{q} \sigma_i X_{t-1} + \rho_t. \tag{2}$$

Where ϑ, σ coefficients of the variables and p, q indicates the delay periods for these variables, ρ is the random error limit.

The delay rank is selected in the ARDL model according to Akaike (AIC) or Schwarz Bayesian Criterion (SBC) before evaluating the OLS model to eliminate sequential or self-correlation of random errors. Pesaran and Shin (2009) recommended a maximum of two periods of deceleration for annual data [6].

In the third stage, the ARDL specification can be obtained for short-term dynamics by building the following error correction model (ECM):

$$\Delta Y_t = c + \sum_{i=1}^{p} \vartheta_i \Delta Y_{t-i} + \sum_{i=0}^{q} \sigma_i \Delta X_{t-i} + \psi ECT_{t-1} + \rho_t \tag{3}$$

Where ECTt-1 is the error correction term, and all the coefficients of the short-term equation are related to the short-term dynamics of the model's proximity to equilibrium, ψ is the coefficient of the error correction term, which measures the adjustment speed at which the imbalance is equal to the corrected in the short-term direction long-run equilibrium.

2 Specifications and model data

The main objective of this paper is measuring the impact of tourism on the economic growth. To do this, we will evaluate this relationship, in particular, by the Tunisia experience (as an example) and, therefore, on the basis of economic theory, as well as empirical models in previous studies (Du & others 2016; Dritsakis 2012) on the same issue [1, 2], equation will be evaluated to measure the impact of tourism on the economic growth (taking into account the exchange rate as an independent variable) in Tunisia during the period: 1985-2017.

$$LGDP_t = \alpha + \beta_1 LTOUR_t + \beta_2 LEXR_t + \epsilon_t \qquad t = 1, 2, \ldots ..., T. \tag{4}$$

Where GDP is real GDP per capita, TOUR is tourism revenues and EXR is the exchange rate.

3 Empirical Results

Before considering the combined cointegration of the ARDL model and evaluation its results, it is important to conduct unit root tests to determine the degree of stability of variables, that is not a necessary condition for using the ARDL model, but the model

does not work exactly if some variables are stationary in the second difference I(2). According to the test results shown in table (1), LGDP & LEX variables stationary at the first difference, with the exception to LTOUR is stationary at the level.

Table 1: Unit root tests

	level series		first difference series	
	PP	*ADF*	*PP*	*ADF*
LGDP	-0.34	-0.31	-5.81*	-5.84*
LTOUR	-3.69*	-2.16	-5.55*	-5.54*
LEXR	0.90	0.75	-5.21*	-5.21*

Notes: data include all variables, they are in natural logarithms.The numbers in parentheses are p-values. * Rejection of null hypothesis at the 1%, 5% and 10% level of significance, respectively. The null hypothesis of these tests is that the time series has a unit root (nonstationary series)
Source: output from Eviews 10.

4 Autoregressive Distributed Lag (ARDL) cointegration technique

The ARDL model, based on the UECM and ARDL Bound Testing Approach models, proposed by Pesaran et al (2001), is most suitable for detecting the cointegration of model variables. Cointegration is estimated by evaluating the UECM model as follows:

$$\Delta LGDP_t = \beta_0 + \sum_{i=1}^{p} \beta_i \Delta LGDP_{t-i} + \sum_{i=0}^{q} \Theta_i \Delta LTOUR_{t-i} + \sum_{i=0}^{q} \omega_i \Delta LEXR_{t-i} + \lambda_1 LGDP_{t-1} + \lambda_2 LTOUR_{t-1} + \lambda_3 LEX_{t-1} + \epsilon_t \quad (5)$$

To test the cointegration of variables, hypotheses are formulated as follows: Null hypothesis (H0): no cointegration : $\lambda_1 = \lambda_2 = \lambda_3 = 0$, Alternative hypothesis (H1): cointegration : $\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq 0$

From the table (2) of cointegration results using the model (ARDL), it is clear that its (F) calculated values are greater than the critical value at a significant 1% level, and then reject the null hypothesis that there is no cointegration between the variables and existence long-term relationships between variables.

Since there is cointegration between model variables, this cointegration includes long-term relationships between these variables, which accept the following formula:

$$LGDP_t = \beta_0 + \sum_{i=1}^{p} \beta_i LGDP_{t-i} + \sum_{i=0}^{q} \Theta_i LTOUR_{t-i} + \sum_{i=0}^{q} \omega_i LEXR_{t-i} + \epsilon_t \quad (6)$$

Table 2: Conditional Error Correction Regression

Regressor	Coefficient	STD
C	1.377743***	0.352211
LGDP(-1)*	-0.223266***	0.052741
LIT(-1)	0.066957***	0.015451
LEX(-1)	0.049234	0.037576
D(LIT)	0.034423*	0.017370
D(LIT(-1))	-0.057722***	0.016469
D(LEX)	-0.017244	0.047454
D(LEX(-1))	-0.101720*	0.052050
D(LEX(-2))	-0.122552**	0.054218
Case 2: Restricted Constant and No Trend		
LIT	0.299900***	0.046453
LEX	0.220519	0.140680
c	6.170858***	0.331246
F-Bounds Test		
F-statistic	20.91202	
K	2	
10%	2.63	3.35
5%	3.1	3.87
2.5%	3.55	4.38
1%	4.13	5

Note: *, ** and *** indicates significance at 10, 5 and 1% level.
Source: output from Eviews 10.

In light of the ARDL criteria (1, 2, 3), to determine the variance of the variables, it was found that the total sum of the balanced weight of the long-term relationship has a negative and significant factor indicating that there is an error correction mechanism, and the equation can be written as follows:

$$LGDP = 0.2999 * LTOUR + 0.2205 * LEXR + 6.1709 \tag{7}$$

The results showed that the error correction parameter was significant by comparing the calculated value of t (t = -4.2332) with the tabulated values of the t_bound test of pesaran et all (2001). Where the test value for the case 2 was -3.66, for two explained estimators and level of significance of 1%. Thus, the error correction parameter is significant at the 1% significance level. The results indicate that the value of this parameter is smaller than zero, which confirms the existence of a cointegration relationship and the possibility of correcting short-term errors to return to

the long-term equilibrium position at a speed of 0.2233 annually, which means that we need 4.29 years to address the effects of one shock occurs in the short term.

On the other hand, the F_bound test indicates that the calculated value is much greater than the tabulated value at the 1% significance level. Which reconfirms the previous result i.e. a common integration relationship between model variables. The results show that the increase in tourism income in Tunisia by 1% leads to an increase in GDP by 0.2999% annually. All are significant values at the level of less than 1% as indicated by Prob. Finally, the results did not show a long-term relationship between the Tunisian dinar exchange rate and GDP.

The Diagnostic tests of the residual of the model. The Breusch-Godfrey Serial Correlation LM test indicates that accept the null hypothesis: that is no autocorrelation, Prob. Chi-Square(2)=0.5530> 0,05. And the Breusch-Pagan-Godfrey indicates that we can't reject the null hypothesis of homoscedasticity, Prob. Chi-Square(8)=0.1884> 0,05. The normality test for Jarque-Bera shows that the null hypothesis cannot be rejected and therefore the residuals are naturally distributed, Probability=0.832278> 0,05

5 The test results on the structural stability of the estimated ARDL modelg

According to Pesaran and Pesaran (1997), the next step after evaluating the formula for the ARDL model is to check the structural stability of short-term and long-term transactions. This means that the data used in this study do not have any structural changes over time. Two tests are used for this: CUSUM (cumulative sum) and CUSUM-sq (CUSUM squared) [7].

The structural stability of the estimated coefficients is achieved if the CUSUM and CUSUMSQ histogram is within critical limits at a significant level of 5%. Consequently, these coefficients are unstable if the graph of the above two tests is exceeded at this level.

From Figure 1, it can be seen that the calculated coefficients are structurally stationary during the study period, the test statistics diagram is within the critical limits at a significant level of 5%.

Conclusions

The purpose of this study is to investigate the relationship between tourism revenues (as an independent variable) and economic growth (as a dependant variable) in Tunisia in the period 1985-2017. To achieve the objective of the study, an autoregressive distributed delay model (ARDL) was used after using the necessary diagnostic tests of time series (i.e. unit root tests).

The results showed a positive impact for tourism on the economic growth in Tunisia during the long term, with elasticity coefficient equal 0.2999, while the ex-

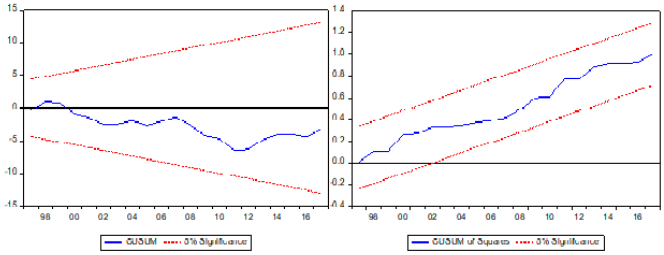


Figure 1: CUSUM and CUSUMSQ
Source: Researchers based on the results (Eviews 10).

change rate was not significant, which might be due to the relative stability in the exchange rate during the period under review.

Therefore, this study recommends focusing on both short and long periods in terms of maintaining moderate exchange rate, and taking appropriate measures (such as marketing campaigns) to attract as many tourists as possible to increase the income from tourism in Tunisia.

References

[1] Dritsakis N. (2012). Tourism development and economic growth in seven Mediterranean countries: a panel data approach. *Tourism Economics*. Vol. **18**, pp. 801-816.

[2] Du D., Ng P., Lew A. A. (2016). Tourism and Economic Growth. *Journal of Travel Research*. Vol. **55**, pp. 454-464.

[3] Holloway, J.C. (1985). *The Business of Tourism. Macdonald and Evans*. Publisher, United Kingdom.

[4] Narayan P. (2005). The saving and investment nexus for China: Evidence from cointegration tests. *Applied Economics*. Vol. **37**, pp. 1979-1990.

[5] Pesaran M. H., Shin Y., Smith R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*. Vol. **16**, pp. 289-326.

[6] Pesaran M., Pesaran B. (2010). *Time Series Econometrics: Using Microfit 5.0.*. Oxford University Press, United Kingdom.

[7] Pesaran M., Pesaran B. (1997). *Working with Microfit 4.0: Interactive Econometric Analysis*. Oxford University Press, United Kingdom.

[8] Poirier R. A. (1995). Tourism and development in Tunisia. *Annals of Tourism Research*. Vol. **22**, pp. 157-171.

[9] Pradhan R., Norman N., Badir Y., Samadhan B. (2013). Transport infrastructure, foreign direct investment, and economic growth interactions in India: The ARDL bounds testing approach. *Social and Behavioral Sciences*. Vol. **104**, pp. 914-921.

[10] Statistics of the Tunisian Ministry of Tourism. *For different years*.

[11] World Tourism Organization (1993). *Recommendations on Tourism Statics*. WTO, Madrid.

Algorithm for regression equation parameters estimation using inverse calculations

Ekaterina B. Gribanova
Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russia
e-mail: katag@yandex.ru

Abstract

The article presents an iteration algorithm for the estimation of regression parameters using inverse operations. Possible usages in the case of linear and nonlinear dependences are provided for. The application of the developed algorithm for solving more complex problems is considered, in particular, results of neural network training using a hybrid algorithm are presented.

Keywords: inverse calculations, regression, optimization.

Introduction

Problems of deriving regression equations have become widespread in various fields: economics, medicine, and technology. Parameter values are determined using a set of input and output data. Both simple methods (for example, the mean-value method) and more complex methods based on optimization problems solving can be used to estimate regression parameters depending on the conditions of the problem, the requirements for the accuracy of the solution and the resources of the researcher. The most common method is the least square method. In the case of linear regression, the following formula can be used to estimate the parameter vector θ:

$$\theta = \left(X^T X\right)^{-1} X^T Y, \tag{1}$$

where X is the matrix of input values, Y is the column vector of output variable values.
In the case of nonlinear dependence, it is necessary to use optimization methods (the method of gradient descent, the Fletcher-Reeves method, etc.) [6].
This paper is devoted to the development of an iteration algorithm based on inverse operations to estimate regression parameters, which can also be used in conjunction with other methods to reduce the search time and improve the solution accuracy.

1 Algorithm for Determining Linear Regression Parameters Based on Inverse Operations

The mathemiatical expectation of remainders shall be equal to zero - it is one of the conditions of the classical linear model of multiple regression. According to the mathematical expectation formula, the condition will be met, if the sum of remainders is close to zero. Fig.1 shows examples of two lines built provided that the sum of

remainders is equal to zero. You can see that line $y = 1.66+0.57x$ is the best in terms of minimizing the sum of squares of deviation remainders. So, the following principle of parameters determination can be considered: derivation of a regression equation so that the sum of remainders is equal to zero, consistent changing of its position while maintaining the equality of the sum of remainders to zero, and selection of the best optimized parameter from the point of view of the value (in our case - the sum of remainder squares).

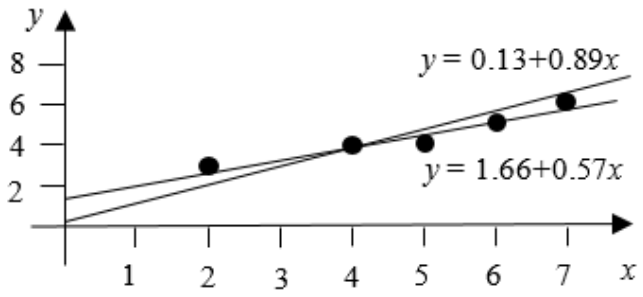


Figure 1: Setting up a regression equation

Let us use a simple example to consider the solution of the problem of determining the parameters of equation $y = b + ax$. The values of x are 5, 6, 2, the corresponding values of y are 30, 50, 10 (the initial values of the parameters $a = 1$, $b = 1$). So, it is necessary to solve the following equation to find the values of the parameters at which the sum of remainders will be zero:

$$3b + 13a = 90 \tag{2}$$

Let us derive parameter a from the equation, the resulting line is shown in Fig.2. It is a set of a and b combinations, at which the equality (2) is fulfilled. Let us now consider two ways of transition from starting point A (1; 1) to the point on the line:

1. Calculate value a by substituting the initial value of b: $a = \frac{90-3\cdot1}{13} = 6.69$. As a result we get point B (1; 6.69).

2. Determine the shortest distance to the line:

$$\Delta a^2 + \Delta b^2 \to min$$

$$(b + \Delta b) + (a + \Delta a)x = 90.$$

The solution of such a problem with the help of inverse operations [5] is considered in [3]. So, it is necessary to solve the system of equations to determine argument increments:

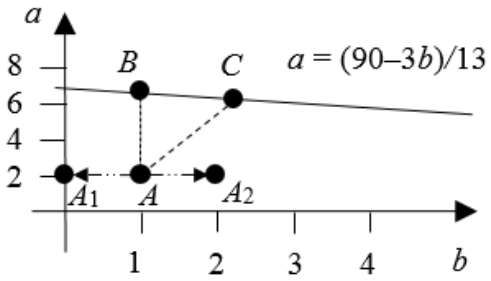


Figure 2: Possible solutions of the equation

$$\begin{cases} 3(1+\Delta b)+13(1+\Delta a)=90 \\ \dfrac{\Delta b}{\Delta a}=\dfrac{3}{13}. \end{cases}$$

So, $\Delta b = 1.247$, $\Delta a = 5.404$. We get point C (2.247; 6.404). Now we need to calculate the value of the sum of error squares at obtained point B (or C), change parameter b by some value to the higher or lower side (points A1, A2), determine the sum of error squares at new points B1, B2 (or C1, C2) and remember the solution with the lowest value of the calculated indicator. The resulting algorithm can be presented as follows (δ- parameter change step, ε - specified accuracy, u - step reduction factor):
Step 1. Initialization of parameters a and b by random numbers ($a^* = a$, $b^* = b$).
Step 2. Calculation of variables to form the sum of errors.
Step 3. Solution of the problem of determining parameters a^* and b^* to achieve the sum of the error equal to zero. Calculation of the sum of squares of remainders:

$$s = \sum_{i=1}^{N}(y_i - d_i)^2$$

where N is the number of observations, y_i is the actual value of the output variable, d_i is the model value of the output variable.
Step 4. Increasing parameter b by the value of δ and determining the parameters to achieve the total error equal to zero (the resulting values of the parameters are equal to a_1, b_1). Calculation of the sum of squares of remainders s_1.
Reduction of parameter b by the value of and determining the parameters to achieve the total error equal to zero (the resulting values of the parameters are equal to a_2, b_2). Calculation of the sum of squares of remainders s_2 .
If $s_1 < s$, then $a^* = a_1, b^* = b_1, s = s_1$,
else if $s_2 < s$, then $a^* = a_2, b^* = b_2, s = s_2$,
otherwise, δ increment reduction in accordance with u factor: $\delta = \frac{\delta}{u}$.
Step 5. Checking the end of the algorithm. If $\delta < \varepsilon$, then the algorithm is completed, otherwise - go to step 4.
In the case of multiple regression, the number of variables changed in step 4 will be equal to the number of variables in the model reduced by 1.
The formula for the formation of the function of the sum of errors is as follows:

$$e = b \cdot N + \sum_{j=1}^{m} \sum_{i=1}^{N} (a_i \cdot x_i^{(j)}) = \sum_{i=1}^{N} y_i,$$

where m is the number of explanatory variables in the model. Then, when using the method based on inverse operations, it is necessary to solve the following system:

$$\begin{cases} N(b + \Delta b) + \sum_{j=1}^{m} \sum_{i=1}^{N} (a_j + \Delta a_j) x_i^{(j)} = \sum_{i=1}^{N} y_i \\ \frac{\Delta b}{\Delta a_j} = \frac{N}{\sum_{i=1}^{N} x_i^{(j)}}, j = 1..m. \end{cases}$$

Changes of weighing factors to achieve the sum of errors equal to zero will be:

$$\Delta b = \frac{bN + \sum_{j=1}^{m} \sum_{i=1}^{N} (a_i \cdot x_i^{(j)}) - \sum_{i=1}^{N} y_i}{-(N + \sum_{j=1}^{m} (\sum_{i=1}^{N} x_i^{(j)})^2 / N)},$$

$$\Delta a_j = \Delta b \frac{\sum_{i=1}^{N} x_i^{(j)}}{N}.$$

Comparing the two ways of calculating the parameters (substitution of the initial value of the parameter in the dependence equation and use of the smallest increments) we can draw the following conclusion. The first way to find a solution is simpler to implement, because it does not require calculating increments by solving a system of equations. However, it is less preferable in those cases where the calculation scheme does not provide for reduction of step δ (step 4 of the algorithm), since it is more likely to find the best solution (move to new point C1 or C2) using the method based on the minimum increments of arguments at each iteration. This is due to the following fact: when using the first method with a slight change in one parameter (step δ is small), the second one can increase or decrease to a much greater extent, which will result in a significant change in the sum of squares of errors, so the best solution in the vicinity of the point being studied may not be found. An example of such a problem is neural network training, when weighing factors are adjusted in each epoch (steps 2-4 of the algorithm, the condition with step δ reduction is eliminated).
Table 1 presents the subsequent iterations when using inverse operations for the example shown in Fig.2. The results obtained using formula (1): b= -10, a = 9.23, the sum of squares of remainders is 61.54.

2 Estimation of Nonlinear Regression Parameters

In the case of nonlinear dependence, if possible, linearization shall be performed. In the case of nonlinear dependence that cannot be linearized, it is necessary to derive one of the variables and calculate partial derivatives of the function obtained to use the approach based on inverse operations. This is how the shortest distance to the

Table 1: Results of the search for linear function parameters

	Step, δ	Parameter b	Parameter a	Sum of error squares, s
1		2.247	6.404	130.791
2	20	-16.742	10.787	82.528
3	10	-7.247	8.596	65.031
4	5	-11.994	9.691	63.374
5	2.5	-9.691	9.143	61.605
6	0.625	-10.214	9.28	61.559
7	0.313	-9.917	9.212	61.542

line tangent to the graph of the line of the level of variables is determined [3]. Now let us consider the following function as an example:

$$y = a \cdot x^b + \mu,$$

where μ is the value of a random remainder.
Initial data: x values are 2, 5, 6, the corresponding values of y are 49, 748, 1297. The initial values of parameters a and b are equal to 1. Here is a corresponding problem of optimization at the first iteration:

$$\Delta a^2 + \Delta b^2 \to min$$

$$(a + \Delta a)2^{(b+\Delta b)} + (a + \Delta a)5^{(b+\Delta b)} + (a + \Delta a)6^{(b+\Delta b)} = 2094$$

Let us consider problem solution using inverse operations. We will derive parameter a: $a(b) = \frac{2094}{2^b+5^b+6^b}$. The value of the derivative at the initial point is $a'(1) = -250.091$. The system of equations is as follows:

$$\begin{cases} \dfrac{\Delta b}{\Delta a} = 250.091 \\ (1 + \Delta a)2^{(1+\Delta b)} + (1 + \Delta a)5^{(1+\Delta b)} + (1 + \Delta a)6^{(1+\Delta b)} = 2094. \end{cases}$$

After solving the system we get: $\Delta a = 0.01215, \Delta b = 3.0382$. After that the derivative value is determined at new point $a'(1 + 0.01215) = -1.745$, and the new system of equations is solved, in which the ratio of increments will be equal to the obtained value. Iterations are performed until the stop condition is met (the change in increments becomes less than the specified accuracy). The results of parameters calculation (accuracy $\varepsilon = 0.001$): a=5.899, b=3.010, s=3.941; to find the root of the equation, Newton's method with an accuracy of 10^-7 was used (the values of parameters obtained using Mathcad package: $a = 5.865$, $b = 3.013$). The time required to solve the problem was 0.0073 seconds (VBA language was used for the implementation). The problem was solved using the method of gradient descent at the descent parameter equal to or less than $5 \cdot 10^-8$, since the objective function increased at

large values. 212,163 iterations and 13.3047 seconds were required to get the solution found with the help of inverse operations using the method of gradient descent. The solution at the given accuracy was found in 0.37 seconds using the random search method with a step of 0.1.

3 Use of the Algorithm for Neural Network Training

The developed algorithm based on inverse operations assumes a step-by-step implementation and can be used in iterative algorithms in conjunction with other methods, for example, the method of gradient descent, as well as in various heuristic algorithms. So, we considered the use of this algorithm when implementing hybrid algorithms for neural networks training. These algorithms imply the use of various methods in training different layers of a neural network. When solving optimization problems, besides hybrid algorithms [4], [2] portfolios of algorithms [4] are created, where every subgroup of observations uses its own algorithm for solving the problem.
Computational experiments were conducted using the developed algorithm (VBA language was used for the implementation). The data set included 2 variables, 100 numbers were randomly generated, output variable values were set depending on the threshold values (Fig.3, the elements belonging to the first class are black, to the second class - gray). The simulation was performed for 50 random implementations, the maximum number of iterations in the stochastic search was 50, while the parameters in the stochastic algorithm were changed in increments of 0.1. The used neural

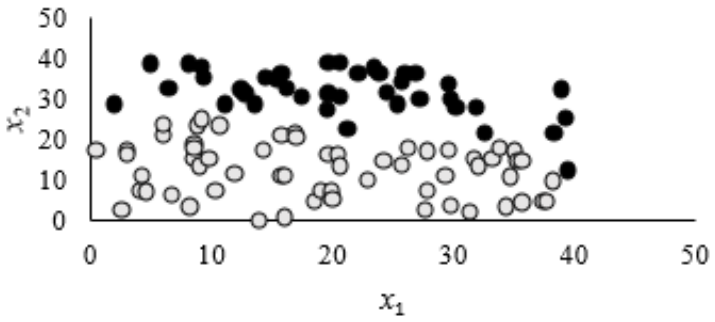


Figure 3: Source data set

network consisting of three layers is shown in Fig.4. The logistic activation function was used for first layer neurons, and the linear function was used for neurons of the subsequent layers. The following options for neural network training were considered:

1. The hybrid algorithm presented in [2], which uses formula (1) to calculate the weighing factors associated with the output neuron.

2. The hybrid algorithm described in [4], which uses the delta rule [7] to calculate the weighing factors associated with the output neuron.

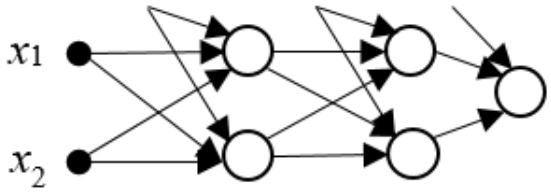


Figure 4: Three-layer neural network

3. A modification of the hybrid algorithm considered in [4], which uses inverse operations to calculate the weighing factors associated with the output neuron (every epoch uses one iteration of the algorithm presented above).

The average value of the sum of squares of errors was obtained for the first algorithm. It was equal to 20.856, the minimum value of the sum of squares of errors was 16.16. Fig. 5 shows values of the sum of squares of errors of the second and third algorithms with the number of epochs equal to 20 and the best values of algorithm parameters (the parameter of the delta rule descent was 0.004, the sum of squares of errors was 17.618, the minimum value of the sum of squares of errors was 10.063; the step for δ parameters change in the method of inverse operations was 0.4, the sum of squares of errors was 13.569, the minimum value of the sum of squares of errors was 6.426). The Figure also shows the results of the algorithm, which is a combination of the delta rule and the method based on inverse operations: in each epoch a solution was determined using these two methods and the one that provided the smallest sum of squares of errors was selected. Using this algorithm the average value of the sum of squares of errors turned out to be 11.64, the minimum value of the sum of squares of errors was 6.852. In this case the solution obtained using the method of inverse operations was adopted as the best one 815 times, the delta rule - 185 times. So, the combination of the two methods provided the smallest value of the sum of squares of errors, however, the time required to solve the problem increased.

Conclusions

The article investigates the possibility of using an algorithm based on inverse operations to solve regression problems. The presented algorithm is simple to implement and allows solving the problems considered in the article at a higher speed as compared to the method of gradient descent and stochastic algorithm. In the course of the work two hybrid algorithms presented in the literature were implemented to train a neural network, and a modification of one of them that consists in using inverse operations to determine the weighing factors associated with the output neuron was proposed. The proposed algorithm can be used to solve regression problems both independently and in conjunction with other methods to increase the speed and

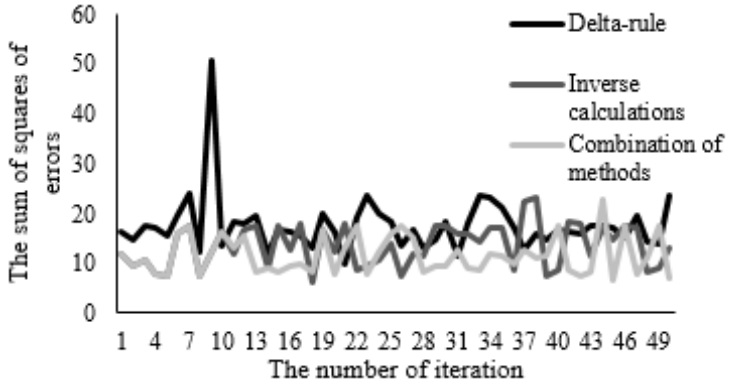


Figure 5: Values of the sum of squares of errors for three algorithms

accuracy of problem solution.

References

[1] Adhikari S.P., Yang S., Slot K., Strzelecki M., Kim H. (2018). Hybrid no-propagation learning for multilayer neural networks. *Neurocomputing*. Vol. **321**, pp. 28-35.

[2] Dudek G. (2018).Generating Random Weights and Biases in Feedforward Neural Networks with Random Hidden Nodes. *Information Sciences*. Vol. **481**, pp. 33-56.

[3] Gribanova E.B. (2018). Methods for Solving Inverse Problems of Economic Analysis by Minimizing Argument Increments. *Proceedings of TUSUR*. Vol. **21**, pp. 95-99.

[4] Odintsov B.Ye. (2004). *Inverse calculations in Making Economic Decisions*. Finance and Statistics, Moscow.

[5] Peng F., Tang K., Chen G., Yao X. (2010). Population-Based Algorithm Portfolios for Numerical Optimization. *IEEE Transaction on evolutionary computation*. Vol. **5**, pp. 782-800.

[6] Tvrdik J., Krivy I. (2004).Comparison of algorithms for nonlinear regression estimates. *Proc. COMPSTAT*, pp. 1917-1924.

[7] Vanneschi L., Castelli M. (2019). Delta Rule and Backpropagation . *Encyclopedia of Bioinformatics and Computational Biology*. Vol. **1**, pp. 621-633.

On rotated versions of one parameter Grubbs's copula

Ludmila K. Shiryaeva
Samara State University of Economics, Samara, Russia
e-mail: Shiryeva_LK@mail.com

Abstract

We consider Grubbs's statistics for normal sample, i.e. the standardized maximum and minimum. One-parameter distribution of these statistics is considered. We extract one-parameter copula by an inversion method from the joint distribution of Grubbs's statistics. We describe properties of the rotated versions of Grubbs's copula. It is proved the existance of domains in which rotated by 90 and 270 degrees Grubbs's copulas coincide with Frechet-Hoeffding upper bound. It is found that rotated by 90 and 270 degrees Grubbs's copulas can model positive dependence between the marginals.

Keywords: symmetric copula, Frechet-Hoeffding lower and upper bounds, rotated copulas, joint distribution function of standardized maximum and minimum.

Introduction

Dependence modeling by means copulas are used in many areas. New copula-function can be extracted from the new joint distribution of random variables. Therefore Grubbs's copula which is extracted from the joint distribution of Grubbs' statistics can be of practical interest. The goal of this article is to investigate the properties of rotated versions Grubbs's copula.

1 On the joint distribution of Grubbs's statistics

Let X_1, X_2, ..., X_{n-1}, X_n be a random sample from a normal distribution with mean a and variance σ^2. We consider one-sided Grubbs' statistics that are extreme studentized deviations of observations from sample mean:

$$T_n^{(1)} = \frac{\max\limits_{1\le i\le n}\{X_i\} - \overline{X}}{S}; \qquad T_{n,(1)} = \frac{\overline{X} - \min\limits_{1\le i\le n}\{X_i\}}{S},$$

where $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the sample mean and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ is the sample variance.

F.E. Grubbs proposed to use these statistics for testing a normal sample on the presence of anomalous observations [3]. It is known that distributions of statistics $T_n^{(1)}$ and $T_{n,(1)}$ coincide, i.e. $P(T_n^{(1)} < t) = P(T_{n,(1)} < t)$.

We denote $F_n^{(1)}(t) = P(T_n^{(1)} < t)$. The distribution function of of Grubbs's statistic can be found with special recursive procedure which has been described in the

book [1] (pp.115-116). A recursive formula for describing the distribution function has the form [1],[8]

$$F_n^{(1)}(t) = P(T_n^{(1)} < t) = \begin{cases} 0,\ t \le \frac{1}{\sqrt{n}},\ n \ge 2; \\ n \int\limits_{\frac{1}{\sqrt{n}}}^{t} F_{n-1}^{(1)}(g_n(x)) f_{T_n}(x)dx,\ \frac{1}{\sqrt{n}} < t \le \frac{n-1}{\sqrt{n}},\ n \ge 3; \\ 1,\ t > \frac{n-1}{\sqrt{n}}, n \ge 2; \end{cases} \quad (1)$$

where

$$f_{T_n}(x) = \frac{1}{n-1}\sqrt{\frac{n}{\pi}}\Gamma\left(\frac{n-1}{2}\right)/\Gamma\left(\frac{n-2}{2}\right)\left(1 - \frac{n}{(n-1)^2}x^2\right)^{\frac{n-4}{2}},\ |x| < \frac{n-1}{\sqrt{n}}; \quad (2)$$

$$\Gamma(x) = \int\limits_0^\infty \xi^{x-1}e^{-\xi}d\xi;$$

$$g_n(x) = \frac{n}{n-1}x \Big/ \sqrt{\frac{n-1}{n-2}\left(1 - \frac{n}{(n-1)^2}x^2\right)},\ |x| < \frac{n-1}{\sqrt{n}},\ n \ge 3. \quad (3)$$

Let $\Lambda_n(t_1, t_2) = P(T_{n,(1)} < t_1, T_n^{(1)} < t_2)$ be the joint distribution function of the statistics $T_n^{(1)}$ and $T_{n,(1)}$. It can be proved that recursive relationships for distribution function $\Lambda_n(\cdot)$ in the case $n > 2$ have the form [5]

$$\Lambda_n(t_1, t_2) = \begin{cases} F_n^{(1)}(t_2),\ \ t_1 \ge \frac{n-1}{\sqrt{n}}; \\ F_n^{(1)}(t_1),\ \ t_2 \ge \frac{n-1}{\sqrt{n}}; \\ n \int\limits_{\frac{1}{\sqrt{n}}}^{t_2} \Lambda_{n-1}\left(\rho_n(t_1, -x), g_n(x)\right) f_{T_n}(x)dx,\ \ (t_1, t_2) \in \Delta_n; \\ 0,\ \ (t_1, t_2) \notin \Delta_n,\ t_1 < \frac{n-1}{\sqrt{n}},\ t_2 < \frac{n-1}{\sqrt{n}}, \end{cases} \quad (4)$$

where distribution function $F_n^{(1)}(t)$ can be calculated with using (1);

$$\rho_n(u, v) = (u + \frac{v}{n-1}) \Big/ \sqrt{\frac{n-1}{n-2}\left(1 - \frac{n}{(n-1)^2}v^2\right)},\ |v| < \frac{n-1}{\sqrt{n}}; \quad (5)$$

functions $g_n(x)$ and $f_{T_n}(x)$ can be calculated with using (3) and (2) correspondingly; $\Delta_n = [1/\sqrt{n} < t_1 < (n-1)/\sqrt{n};\ 1/\sqrt{n} < t_2 < (n-1)/\sqrt{n}\,]$.

In case $n = 2$ they are given by

$$\Lambda_2(t_1, t_2) = \begin{cases} 1,\ (t_1, t_2) \in \Delta_2,\ \Delta_2 = [\frac{\sqrt{2}}{2} < t_1 < \infty; \frac{\sqrt{2}}{2} < t_2 < \infty]; \\ 0, \qquad\qquad (t_1, t_2) \notin \Delta_2. \end{cases} \quad (6)$$

Note some properties of the joint distribution function $\Lambda_n(t_1, t_2)$ which can be derived from (4) .

1. Function $\Lambda_n(t_1, t_2)$ is symmetrical [5], i.e.

$$\Lambda_n(t_1, t_2) = \Lambda_n(t_2, t_1).$$

2. It is valid for $(t_1,\, t_2) \in \Sigma_n$ and $n \geq 3$ [7]:

$$\Lambda_n(t_1, t_2) = F_n^{(1)}(t_1) + F_n^{(1)}(t_2) - 1, \tag{7}$$

where $\Sigma_n = [\frac{1}{\sqrt{n}} \leq t_1 < \frac{n-1}{\sqrt{n}}; \theta_n(t_1) \leq t_2 < \frac{n-1}{\sqrt{n}}]$;

$$\theta_n(t_1) = \frac{t_1}{n-1} + \sqrt{n-2}\sqrt{1 - \frac{nt_1^2}{(n-1)^2}}. \tag{8}$$

3. In the case $n = 3$ we can write[6]

$$\Lambda_3(t_1, t_2) = \begin{cases} 0, & (t_1, t_2) \in \Delta_3 \backslash \Sigma_3 \\ F_3^{(1)}(t_1) + F_3^{(1)}(t_2) - 1, & (t_1, t_2) \in \Sigma_3, \end{cases} \tag{9}$$

where $F_3^{(1)}(t) = \frac{3}{\pi} \arcsin\left(\frac{\sqrt{3}}{2}t\right) - \frac{1}{2}$, $1/\sqrt{3} \leq t \leq 2/\sqrt{3}$.

2 Construction of Grubbs copula

To construct Grubbs's copula by extruction from Λ_n we apply Sclar's Theorem [4].

Denote $\phi_n(x)$ is inverse- function for $F_{n,(1)}$, i.e. the equalities are true

$$F_n^{(1)}(\phi_n(x)) = x, \ \forall x \in [0, 1],$$

and

$$\phi_n(F_n^{(1)}(t)) = t, \ \ \forall t \in [1/\sqrt{n}, (n-1)/\sqrt{n}].$$

Then Grubbs's copula $C^{Gr} : [0, 1]^2 \to [0, 1]$ has the following form

$$C^{Gr}(u, v; n) = \Lambda_n(\phi_n(u), \phi_n(v)) \tag{10}$$

Note some properties of the Grubbs's copula which can be deducted from the properties of the joint distribution function $\Lambda_n(\cdot)$.

1. Grubbs's copula is symmetrical [6], i.e.

$$C^{Gr}(u, v; n) = C^{Gr}(v, u; n), \ \ \forall (u, v) \in [0, 1]^2.$$

2. Let $n > 3$ and $\Xi_n = [0 \leq u \leq 1; \delta_n(u) \leq v \leq 1]$. Then $\forall (u, v) \in \Xi_n$ Grubbs's copula coincides with Frechet-Hoeffding lower bound [7], i.e.

$$C^{Gr}(u, v; n) = u + v - 1, \tag{11}$$

where

$$\delta_n(u) = F_n^{(1)}(\theta_n(\phi_n(u)); \tag{12}$$

functions $\theta_n(\cdot)$ and $F_n^{(1)}(\cdot)$ is calculated by formulas (8) and (1), respectively.
3. In the case $n = 3$ Grubbs's copula coincides with Frechet-Hoeffding lower bound [6], i.e.

$$C^{Gr}(u, v; 3) = \max(u + v - 1; 0), \ \forall (u, v) \in [0, 1]^2. \tag{13}$$

Analyzing properties 2—3 we lead to conclusion that the bound $\delta_n(u)$ of the domain Ξ_n contains the points with coordinates $(0, 1)$ and $(1, 0)$. Besides, $\delta_3(u) = 1 - u$. If $v = \delta_n(u)$ then $\forall u \in [0, 1]$ and $n \geq 3$ we obtain

$$C^{Gr}(u, \delta_n(u); n) = u + \delta_n(u) - 1.$$

Hence, $\forall u \in [0, 1]$ and $n \geq 3$ we have

$$C^{Gr}(u, \delta_n(u); n) = \delta_n(u) - \delta_3(u).$$

We can write in accordance with copulas definition [4]

$$C^{Gr}(u, v; n) \geq 0, \quad \forall (u, v) \in [0, 1]^2, \quad n \geq 3.$$

Then

$$\delta_n(u) \geq \delta_3(u), \quad \forall u \in [0, 1], \quad n \geq 3. \tag{14}$$

Thus, the domain Ξ_n is bounded by the lines $u = 1$, $v = 1$ and the curve $v = \delta_n(u)$. With the increasing parameter n the bound $v = \delta_n(u)$ is removed from the main diagonal $u + v = 1$ of the unit square.

3 Rotated versions of Grubbs's copula

Grubbs's copula allows to describe negative dependencies between marginals. In addition to this copula we can introduce its rotated versions. Rotation by 180 degrees leads to the survival copula. Survival copula C_{180}^{Gr} can be extracted from the corresponding joint survival function $\overline{\Lambda}_n(t_1, t_2) = P(T_{n,(1)} > t_1, T_n^{(1)} > t_2)$. Then we have [4]

$$\overline{\Lambda}_n(t_1, t_2) = C_{180}^{Gr}(\overline{F_n^{(1)}}(t_1), \overline{F_n^{(1)}}(t_2); n),$$

where $\overline{F_n^{(1)}}(t) = P(T_n^{(1)} > t) = P(T_{n,(1)} > t)$.

It is valid the following equality

$$C_{180}^{Gr}(u, v; n) = u + v - 1 + C^{Gr}(1 - u, 1 - v; n). \tag{15}$$

When rotating Grubbs's copula by 90 and 270 degrees we obtain its rotated versions which allow to describe the positive dependence.

Rotated versions of C^{Gr} are given as follows [2]

$$C_{90}^{Gr}(u, v; n) = v - C^{Gr}(1 - u, v; n); \tag{16}$$

$$C_{270}^{Gr}(u, v; n) = u - C^{Gr}(u, 1 - v; n). \tag{17}$$

Some properties of rotated versions of Grubbs's copula are contained in the next theorems.

Theorem 1. *Let* $\Xi_n^{(90)} = \{0 \le u \le 1; \delta_n(1-u) \le v \le 1\}$, $\Xi_n^{(180)} = \{0 \le u \le 1; 0 \le v \le 1 - \delta_n(1-u)\}$, $\Xi_n^{(270)} = \{0 \le u \le 1; 0 \le v \le 1 - \delta_n(u)\}$ *and function* $\delta_n(\cdot)$ *is defined in accordance with (12). Then for* $n \ge 3$ *the next equalities take place*

$$C_{90}^{Gr}(u,v;n) = \min(u,v) \equiv u, \ \ \forall (u,v) \in \Xi_n^{(90)}. \tag{18}$$

$$C_{180}^{Gr}(u,v;n) = 0, \ \ \forall (u,v) \in \Xi_n^{(180)}. \tag{19}$$

$$C_{270}^{Gr}(u,v;n) = \min(u,v) \equiv v, \ \ \forall (u,v) \in \Xi_n^{(270)}. \tag{20}$$

Proof. Taking into account formula (11) we obtain

$$C^{Gr}(1-u,v;n) = v - u, \ \ \forall (u,v) \in \Xi_n^{(90)}.$$

Applying (16) we can write

$$C_{90}^{Gr}(u,v;n) = u, \ \ \forall (u,v) \in \Xi_n^{(90)}. \tag{21}$$

Taking into account inequality (14) we have $\delta_n(1-u) \ge u, \ \forall u \in [0;1]$. Therefore $\forall (u,v) \in \Xi_n^{(90)}$ we can write $u \le \delta_n(1-u) \le v$. Hence, $\forall (u,v) \in \Xi_n^{(90)}$ we obtain $\min(u,v) = u$. Then the formula (21) takes the form (18). Similarly, it is possible to prove the validity of equality (19) and (20).

Theorem 2. *Let* $n = 3$. *Then rotated copulas* C_{90}^{Gr} *and* C_{270}^{Gr} *coincide with Frechet-Hoeffding upper bound and* C_{180}^{Gr} *coincides with Frechet-Hoeffding lower bound, i.e.*

$$C_{90}^{Gr}(u,v;n) = C_{270}^{Gr}(u,v;n) = \min(u,v), \ \ \forall (u,v) \in [0;1]^2; \tag{22}$$

$$C_{180}^{Gr}(u,v;n) = \max(u+v-1;0), \ \ \forall (u,v) \in [0;1]^2. \tag{23}$$

Proof. If $n = 3$ then $\delta_3(u) = (1-u)$. Taking into account formula (13) we can write

$$C^{Gr}(u,v;3) = \begin{cases} 0, & 0 \le u \le 1, 0 \le v \le 1-u \\ u+v-1, & 0 \le u \le 1, 1-u \le v \le u, \end{cases}$$

Hence,

$$C^{Gr}(1-u,v;3) = \begin{cases} 0, & 0 \le u \le 1, 0 \le v \le u \\ v-u, & 0 \le u \le 1, u \le v \le 1, \end{cases}$$

Taking into account formula (16) we obtain (22). Similarly, it is possible to prove the validity of equality (23).

Conclusions

1. While Grubbs's copula models negative dependence between the marginals, its rotated by 90 and 270 degrees versions can model positive dependence between the marginals.
2. For all parameter's values $n \ge 3$ there are domains in which rotated by 90 and 270 degrees Grubbs's copulas coincide with Frechet-Hoeffding upper bound.
3. If copulas parameter $n = 3$ then rotated by 90 and 270 degrees Grubbs's copulas completely coincide with Frechet-Hoeffding upper bound and rotated by 180 degrees Grubbs's copula completely coincides with Frechet-Hoeffding lower bound.

References

[1] Barnett V., Lewis T. (1984).*Outliers in statistical data*. John Wiley Sons, Chichester.

[2] Brechmann E., Schepsmeier U. (2013). Modeling dependence with C- and D-Vine Copulas: The R package CDVine. *Joirnal of Statistical Software*. Vol. **52**. No. 3.

[3] Grubbs F.E. (1950). Sample Criteria for Testing Outlying observations. *Ann. Math. Statist*. Vol. **21**, pp. 27–58.

[4] Nelsen R.B. (2006).*An Introduction to Copulas*. Springer-Verlag, New York.

[5] Shiryaeva L.K. (2014). On null and alternative distribution of statistics of two-side discordancy test for an extreme outlier. *Russian Math.(Iz. VUZ)*. Vol. **58:10**, pp. 52–66.

[6] Shiryaeva L.K. (2015). On tail dependence for Grubbs' copula-function. *Russian Math.(Iz. VUZ)*. Vol. **59:12**, pp. 27–58.

[7] Shiryaeva L.K. (2019). On three-parameter for Grubbs' copula-function. *Iz. VUZ. Matem*. Vol. **3**, pp. 38–55.

[8] Zhang J. , Keming Y. (2006). The null distribution of the likelihood-ratio test for one or two outliers in a normal sample. *Sociedad de Estatistica e Investigacion Operativa Test*. Vol. **15**. No. 1, pp. 141–150.

Logistic regression model of student retention based on analysis of the Bolasso regularization path

Anastasiia Yu. Timofeeva and Alena A. Borisova
Novosibirsk State Technical University, Novosibirsk, Russia
e-mail: a.timofeeva@corp.nstu.ru, bborisova2012@yandex.ru

Abstract

For exploring student characteristics of retention, a logistic regression model is constructed. Candidate predictors correlate strongly. Hence the problem of variable selection arises. For stabilizing the LASSO the Bolasso is used. But the choice of the regularization parameter based on cross-validation does not give a well-interpretable model. Therefore, approach to analyzing the regularization path is being developed to select a model structure that includes only relevant covariates. For this, simple indicators of multicollinearity and significance are introduced. The applicability of the proposed approach is shown by example of identifying the reasons why students are left to study in the master's program.

Keywords: LASSO, bootstrap, regularization path, student retention, logistic regression, variable selection.

Introduction

Logistic regression is often used to estimate the impact of various factors on student retention [7]. From a large number of attributes related to university drop-out and persistence, the LASSO regression selects the relevant variables during the model construction process. The LASSO estimates are however known to be highly unstable for several reasons. First, irrelevant attributes can enter the model randomly with strictly positive probability. This problem can be solved using the Bolasso [1] that asymptotically selects with overwhelming probability the correct relevant variables. Secondly, the regression solution is sensitive to the penalty parameter chosen. Cross-validation is often used to select an optimal value of the regularization parameter. For stabilizing the LASSO against cross-validation variability the percentile-lasso is introduced [6]. But cross-validation evaluates a model prediction performance, so too many covariates can be selected, and the model can lack interpretability. A "one-standard error" rule [4] gives the more parsimonious model, but not always a sufficiently sparse solution. Third, when variables are highly correlated, a single variable is picked at random. Some kinds of modified penalty functions [3] have been developed to this instability. However, they lose some of computationally attractiveness.

In contrast to the loss function modification, we propose to analyze the Bolasso regularization path. This paper introduces new simple indicators that allow to find a meaningful model structure. The model selection procedure is semi-automatic and incorporates a subjective assessment of strong correlated input variables. This is

a more flexible approach, since the initial set of variables can be extended by interaction effects and combined categorical variables. The approach is however applicable only in medium-dimensional problems (no more than a hundred variables). It is this number of factors that influence the student retention that can be extracted from databases of educational institutions.

1 Logistic regression

Let us introduce a Bernoulli random variable η which takes the value 1 in case of student retention. Suppose a conditional probability of student persistence for a given vector $\mathbf{x}$ of explanatory variables is modeled by logistic regression [5]:

$$\Pr(\eta = 1|\mathbf{x}) = \mathbb{E}(\eta|\mathbf{x}) = F(\mu + \mathbf{x}'\boldsymbol{\theta}), \tag{1}$$

where F is a logistic distribution function, $\boldsymbol{\theta}$ is a vector of effects of explanatory variables to be estimated, μ is an intercept.

The regression coefficients in (1) can be estimated by minimizing the negative log-likelihood:

$$-L(\mu, \boldsymbol{\theta}) = \sum_{i=1}^{N} \rho_{\mu,\boldsymbol{\theta}}(\mathbf{x}_i, y_i), \tag{2}$$

where y_i is a value of η for i-th student, $\mathbf{x}_i$ is a vector of values of covariates for i-th student, N is a total number of students, the loss function is defined as $\rho_{\mu,\boldsymbol{\theta}}(\mathbf{x}_i, y_i) = -y_i(\mu + \mathbf{x}_i'\boldsymbol{\theta}) + \log(1 + \exp(\mu + \mathbf{x}_i'\boldsymbol{\theta}))$.

The initial set of variables usually includes some irrelevant attributes and redundant ones. This makes the model too complex. Thus, optimization of (2) results in overfitting that can produce misleading regression coefficients. The regularization can be performed in order to enhance the interpretability of the logistic model.

2 LASSO regularization

The LASSO regularization imposes a constraint on the model parameters to shrink some regression estimates towards zero [4]. The estimation problem is defined as follows:

$$\min_{\mu,\boldsymbol{\theta}} \frac{1}{N} \sum_{i=1}^{N} \rho_{\mu,\boldsymbol{\theta}}(\mathbf{x}_i, y_i) + \lambda \|\boldsymbol{\theta}\|_1, \tag{3}$$

where λ is a regularization parameter, $\|\cdot\|_1$ is the ℓ_1-norm.

The sparsity of the solution in the problem (3) depends on the choice of the regularization parameter λ.

2.1 Bolasso

Drastic changes in LASSO-regression coefficients are however known to arise in the case of small data perturbations. So for a given value of λ the subset of selected

variables is also unstable. This can be revealed by resampling methods such as the bootstrap. The intersection of subsets of selected variables for the bootstrap samples defines the Bolasso model structure [1]. Thus, the final subset contains only those attributes that were entered simultaneously in all LASSO regression models. As a result, the Bolasso can select the correct relevant variables.

The Bolasso regression estimates are computed simultaneously for a large number of regularization parameters. That allows to find the entire regularization path. Let a decreasing sequence of the regularization parameter values be given $\lambda_1, ..., \lambda_T$, that is, λ_1 corresponds to the simplest model with fewer variables.

The Bolasso is sometimes too strict in intersecting models. We preferred a soft version of the Bolasso (referred to as Bolasso-S), where we select those variables which are present in at least 90% of the bootstrap replications. Then, at each λ_t, the empirical frequency f_{jt} of that the estimates of j-th covariate do not get shrunk to exactly zero is calculated. The subset of relevant variables is defined by the condition $f_{jt} \geq \gamma$, where γ is the threshold value (1 for the Bolasso, 0.9 for the Bolasso-S). For each such subset, the parameter vector of the ordinary logistic regression is estimated by minimizing (2). The vectors obtained may have a different number of elements depending on t. In order to avoid this, we fill the missing elements of the vectors (the coefficients of eliminated variables) with zeros. Denote such a vector as $\boldsymbol{\theta}_t$.

The optimal value of the regularization parameter and the corresponding model structure can be selected using cross-validation. However, it provides a good predictive ability of the model, but does not guarantee its simplicity and interpretability. Moreover, in presence of strong correlations between variables, the estimates cannot be consistent. Further, we propose to analyze the entire regularization path in order to select an interpretable model structure.

2.2 Analysis of the Bolasso Regularization Path

There are two main problems that arise by entering a new variable in the model:

- this variable has no significant effect on the response and is not correlated with other input variables;
- this variable is significantly related to the response and other relevant covariates already entered in the model.

The significance of attribute effect can be tested by t-statistic for the corresponding coefficient. If the calculated p-value will be above the threshold α chosen for statistical significance (usually the 0.10, the 0.05, or 0.01 level), then the effect of this attribute on the response is insignificant.

However, in the second case, the t-test can give misleading results due to the multicollinearity. By adding some predictor variables to a regression model that are highly related, the estimated regression coefficients change drastically. It is proposed to detect this case by calculating the maximum absolute deviation of the parameter

estimates:

$$D_t = \max_{i \in I_t} \left| \frac{\theta_{it} - \theta_{i(t-1)}}{\theta_{i(t-1)}} \right|, \tag{4}$$

where $I_t = \{i : \theta_{it} \neq 0 \& \theta_{i(t-1)} \neq 0\}$, θ_{it} is i-th element of the vector $\boldsymbol{\theta}_t$ corresponding to the value of the regularization parameter λ_t.

If the deviation (4) is large, then multicollinearity may be present in a model. In this case, strong correlated input variables are evaluated subjectively. The decision on which of the correlated variables should be added to the model can be made on the basis of their correlation with the response or economic considerations.

Additionally, it makes sense to check whether a large deviation is actually caused by multicollinearity, by calculating the maximum correlation between the predictors. Since student retention is influenced by numerical as well as nominal attributes, it is proposed to use mutual information of the variables X_i, X_j to measure the dependence between them.

$$MI(X_i, X_j) = H(X_i) - H(X_i|X_j),$$

calculated on the basis of entropies:

$$H(X_i) = -\sum_{k=1}^{K} p(X_i = k) \log p(X_i = k),$$
$$H(X_i|X_j) = -\sum_{k=1}^{K} \sum_{m=1}^{M} p(X_i = k, X_j = m) \log p(X_i = k|X_j = m),$$

where $p(A)$ is the relative frequency of the event A, $p(X_i = k|X_j = m)$ is the relative frequency with which i-th variable takes the value k under the condition that j-th variable is equal to m, K, M are the numbers of categories of the i-th and j-th variables, respectively. The discretization of the numerical attributes is performed. In empirical study quartiles were used as discretization thresholds.

The multicollinearity indicator proposed is the maximum of the mutual information of all pairs of predictors entered in the model for a given λ_t:

$$\max MI_t = \max_{i,j \in U_t} MI(X_i, X_j), \tag{5}$$

where $U_t = \{i, j : f_{it} \geq \gamma \& f_{jt} \geq \gamma \& i \neq j\}$.

The jump of the indicator (5) should detect that the model structure at λ_t differs from one at λ_{t-1} in that a predictor correlated with the other variables was included.

The proposed analysis of the the regularization path involves the following steps.

Step 1. For all values of the regularization parameter, the multicollinearity indicators $\log D_t, \max MI_t$, $t = 1, ..., T$ are calculated.

Step 2. Student's t-test is used to determine which regression parameters are not significantly different from zero. The p-value is computed for such parameters which are added to the model at λ_t ($\theta_{i(t-1)} = 0 \& \theta_{it} \neq 0$) or excluded from the model at λ_t ($\theta_{i(t-1)} \neq 0 \& \theta_{it} = 0$). Denote the set of p-values calculated for a given λ_t as PV_t.

Step 3. Set $t := 2$.

Step 4. If $\log D_t < 1$ and $|\max MI_t - \max MI_{t-1}| < \delta$, where δ is a small positive value, then go to step 5, otherwise, multicollinearity is detected when the variable

subset at λ_{t-1} changes to the subset at λ_t. Then the redundant (correlated or duplicated) variables are eliminated from the model or interaction effects (or combined categorical variables) are added to the model.

Step 5. Exclude from the model those covariates for which $PV_t > \alpha$, except for those variables that were not significant due to multicollinearity and were left in the model at step 4.

Step 6. Set $t := t + 1$ and go to step 4, while $t \leq T$.

We have implemented the proposed analyze the regularization path using a standard set of `R` packages. In addition a publicly available package `glmnet` [2] is used for fitting the entire lasso regularization path for logistic regression model. For the Bolasso, samples of size N were drawn uniformly at random with replacement from the original data. Three hundred bootstrap replications were generated.

3 Empirical results

The empirical study is devoted to analysis of choice of a master's degree at the university. The dataset is collected from the information system of the largest university in Novosibirsk. Table 1 presents the original set of variables influencing the choice of a master's degree. For categorical attributes the first level is the reference. The set contains strong correlated predictors, in particular Grade Point Average (GPA) on a 5-point system and on a 100-point one.

Table 1: Candidate predictors of master's degree applicant's choice

Index	Variable	Categories
1	Faculty	8 technical, 4 humanitarian
2	Year of the first publication	No, 1st-2nd, 3rd, 4th
3-5	Number of publications: total, 3rd, 4th course	0, 1, 2, more than 2
6	WoS publication	No, yes
7	Publication language	No or russian, foreign
8	Independence of publications	No, independent, co-authored
9-12	GPA on 1st-4th years	On a 5-point system
13-16	GPA on 1st-4th years	On a 100-point system
17	Government grant support	Budget, contract
18	Residence	other, Novosibirsk
19	Year of the first research	No, 1st-2nd, 3rd, 4th
20	Obtaining funding	No, yes
21	Amount of funding	in logarithms, "no" replaced by 0

A preliminary analysis revealed that the probability of choosing a master's degree differs for technical and humanitarian faculties. Therefore, it was decided to build two different logistic models. Figure 1 shows the empirical frequencies of selecting

any given variable for the Bolasso as the regularization parameter varies. The large frequencies are in white, the small values are in black. The horizontal axis shows the values of $-\log \lambda_t$.

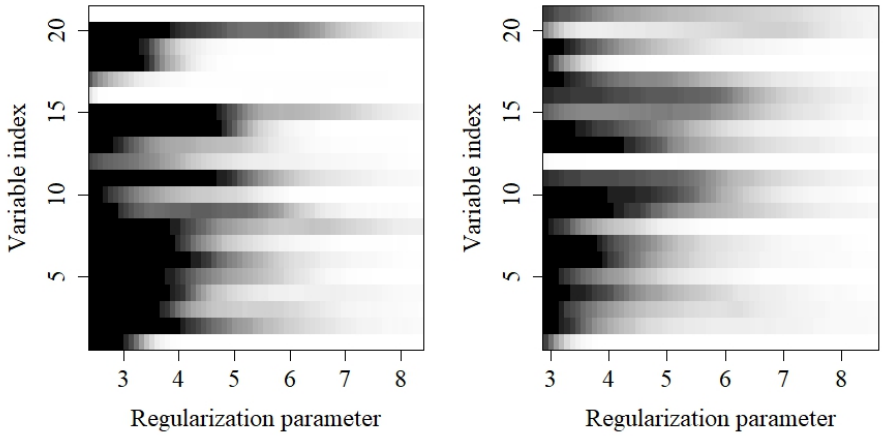


Figure 1: The relative frequencies of variable selection
for technical (left) and humanitarian (right) faculties

By estimating a model for technical faculties, the Bolasso does not work so badly. First variable included in the model is the amount of funding. Further, significant factors are added: GPA on 4th year on a 100-point system, the government grant support, the faculty, the residence, and the year of the first research. Next, an insignificant variable of GPA on 2th year on a 5-point system is included. Adding GPA on 2th year on a 100-point system causes multicollinearity. In Figure 2 this can be clearly seen from the peak value of $\log D_t$ with a sharp increase in $\max MI_t$. At the same value of λ_t, an insignificant factor of the publication language is added. Further, the number of publications on the 4th year and their total number are added. They are significant, but highly correlated, so we enter into the model the second only. All further included variables are insignificant or correlated (duplicated) and ignored.

Figure 3 shows that for humanitarian faculties at large values of the regularization parameter, only the first three changes in the model structure led to selection of significant attributes: GPA on the 4th year on a 5-point system, the faculty, the residence, the independence of publications. The number of publications on the 4th year is significant at 10% level, we will neglect it. Further, only insignificant variables are added while the total number of publications and the number of publications for the 3rd year are included. This causes a jump in $\log D_t$. $\max MI_t$ also increases sharply. Thus, only the total number of publications should be included in the model. However, this attribute has the same level with the independence of publications, namely "no publications". Therefore, it was decided to combine these two categorical variables into one with the reference level "no publications".

The significant factors added to the model at large values of $-\log \lambda_t$ are the government grant support, the obtaining and amount of funding. They correlate

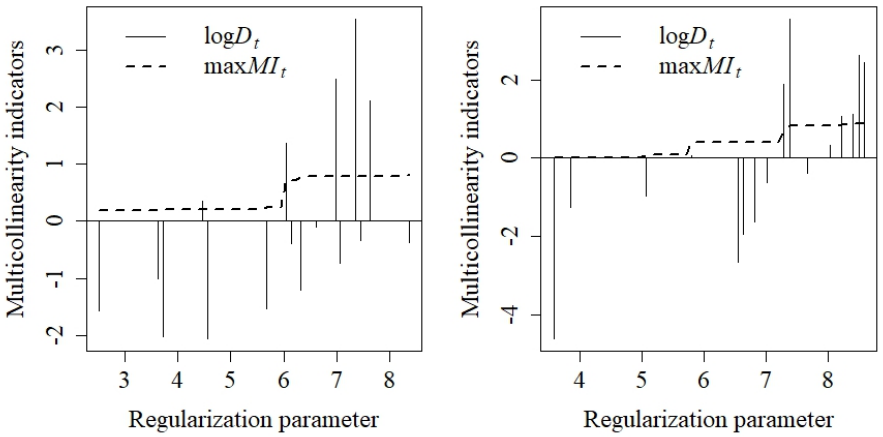


Figure 2: The multicollinearity indicators
for technical (left) and humanitarian (right) faculties

strongly, this causes the last two peaks of $\log D_t$. All budget students are supported by a government grant. So it is impossible for them to estimate the impact of not funding. At the same time, for contract students, funding is so rare that it is difficult to estimate the impact of its amount. Therefore, it was decided to include the factor Contract and the interaction between Contract and Obtaining funding. Thus, in order to obtain the final structure of the model, it was necessary to analyze the entire regularization path up to the minimum value of the regularization parameter.

Conclusions

The Bolasso is used to stabilize the results of the LASSO regression estimation. However, the choice of the regularization parameter that provides a good interpretation can be a serious issue. In the presence of multicollinearity the parameter estimates may get counterintuitive signs. An empirical study showed that in such cases it is impossible to find the optimal value of the regularization parameter, which ensures the simplicity and meaningfulness of the regression model. Therefore, the proposed analysis of the regularization path is a good solution to the problem. It allowed us to understand what predictors it makes sense to include in the model. As a result, we extended the original set of attributes by interaction effects and combined categorical variables. This would not have been achieved by optimal choice of the regularization parameter.

Acknowledgements

The reported study was funded by Russian Ministry of Education and Science, according to the research project No. 2.2327.2017/4.6.

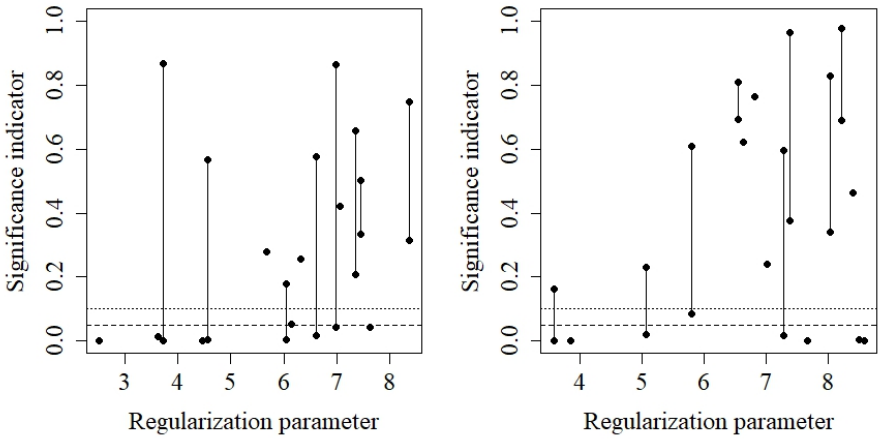


Figure 3: The significance indicator
for technical (left) and humanitarian (right) faculties

References

[1] Bach F.R. (2008). Bolasso: model consistent lasso estimation through the bootstrap. *Proceedings of the 25th international conference on Machine learning*. ACM, pp. 33-40.

[2] Friedman J., Hastie T., Tibshirani R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*. Vol. **33**, pp. 1-22.

[3] Grave E., Obozinski G.R., Bach F.R. (2011). Trace lasso: a trace norm regularization for correlated designs. *Advances in Neural Information Processing Systems*. Curran Associates, Inc., pp. 2187-2195.

[4] Hastie T., Tibshirani R., Wainwright M. (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations*. CRC Press.

[5] Hosmer Jr.D.W., Lemeshow S., Sturdivant R.X. (2013). *Applied logistic regression*. John Wiley & Sons.

[6] Roberts S., Nowak G. (2014). Stabilizing the lasso against cross-validation variability. *Computational Statistics & Data Analysis*. Vol. **70**, pp. 198-211.

[7] Wohlgemuth D., et al. (2007). Financial, academic, and environmental influences on the retention and graduation of students. *Journal of College Student Retention: Research, Theory & Practice*. Vol. **8**, pp. 457-475.

Analysis of the methods of the Kriging family and GWR for transport speeds prediction models development

V. S. Timofeev, A. S. Veselova, K. V. Teselkina
Novosibirsk State Technical University, Novosibirsk, Russia
e-mail: v.timofeev@corp.nstu.ru, veselova8nastya@gmail.com

Abstract

The article studies geostatistical methods for estimating individual transport speeds in a populated area. The authors proposed spatial models of velocities, namely the model of geographically-weighted regression and and interpolation using universal kriging and kriging with external drift. Optimal parameters of the models were chosen and a comparison of the proposed methods was made. The study was carried out on the data of road users in Novosibirsk, Russia.

Keywords: traffic speeds, spatial speed model, geographically weighted regression model, universal kriging, kriging with external drift.

Introduction

Residents of the modern metropolis constantly face the problem of traffic congestion and inefficient operation of the transport system in general. The authorities of population center need tools to support management decisions to optimize the transportation system. The model of the transport system can serve as an instrument that displays not only the current state of transport complex, but also predicts the consequences of management impacts. Practically suitable methods of creating such models are extremely limited nowadays.

With the development of spatial data collection technologies, new methods of statistical analysis that take into account the location of the objects of research appear. Such methods include the method of geographically-weighted regression (GWR) [2] and methods of the kriging family [5]. In the paper, it is proposed to adapt these methods for constructing a model of transport speed in a populated area that is part of the transport system model as a whole. Using the methods of geostatistics allows to take into account the spatial heterogeneity of the data and to obtain more accurate models. Application of the methods of geographically weighted regression and universal kriging to estimate transport speeds was proposed for the first time in [1]. Interpolation of transport speeds by kriging with external drift and comparison of the quality of estimation of all three methods in this paper was carried out for the first time.

This research has been supported by the Ministry of Education and Science of the Russian Federation as part of the state task (project No 2.7996.2017).

1 Geographically weighted regression

We divide the entire study area into m areas, each characterizes a certain point of interest with coordinates (u_i, v_i). There are n_i observation in each such area. Total data used for analysis contains N observations. Then the model of geographically-weighted speed regression is following:

$$y^i = \beta_0(u_i, v_i) + \sum_{k=1}^{r} \beta_k(u_i, v_i) x_k^i + \epsilon^i, i = 1...m, \quad (1)$$

where y^i - individual transport speed (km/h) in some area which characterizes by point i with coordinates (u_i, v_i); $\beta_k(u_i, v_i)$ - an estimated parameter that is an implementation of a continuous function $\beta_k(u, v)$ at point i; ϵ^i - random error at point i , $\forall i \neq j$ $cov(\epsilon^i, \epsilon^j) = 0$; x_k^i - significance of the explanatory factor F_k. The paper considers $r = 16$ input factors; the list of them and their possible values are:

- F_1 - Road section type, $F_1 \in$ {rural road, highway with signal controled traffic, main street of regional, importance (pedestrian-transport), main street of district importance (transport-pedestrian), non, stop city-wide main street , city-wide main street with signal controled traffic and transit flow, city-wide main street with signal controled traffic, street of urban housing, street in industrial district, pedestrian street, street with a dedicated lane};
- F_2 - Number of lanes, $F_2 \in$ {1,2,3,4,5,6 };
- F_3 - The maximum allowed speed (km/h), $F_3 > 0$;
- F_4 - Max. throughput (vel./h), $F_4 > 0$;
- F_5 - Intersection regulation of the road section start, $F_5 \in$ {1-traffic light, 0-without regulation};
- F_6 - Intersection regulation of the road section end, $F_6 \in$ {1-traffic light, 0-without regulation};
- F_7 - Right turn at the road section start, $F_7 \in$ {1- allowed, 0-denied };
- F_8 - Left turn at the road section start, $F_8 \in$ {1- allowed, 0-denied };
- F_9 - Backward turn at the road section start, $F_9 \in$ {1- allowed, 0-denied };
- F_{10} - Go straight at the road section start, $F_{10} \in$ {1- allowed, 0-denied };
- F_{11} - Right turn at the road section end, $F_{11} \in$ {1- allowed, 0-denied };
- F_{12} - Left turn at the road section end, $F_{12} \in$ {1- allowed, 0-denied };
- F_{13} - Backward turn at the road section end, $F_{13} \in$ {1- allowed, 0-denied };
- F_{14} - Go straight at the road section end, $F_{14} \in$ {1- allowed, 0-denied };

- F_{15} - Share of the traversed path to the road section end, $F_{15} \in [0,1]$;
- F_{16} - Length of the road section (km), $F_{16} > 0$.

Some factors are qualitative and have several possible levels, therefore, the corresponding sets of dummy variables [4] were used to estimate the regression parameters. The parameters of the regression equation (1) at each point of interest i with coordinates (u_i, v_i) can be estimated by the following formula:

$$\hat{\beta}(u_i, v_i) = (X_i^T W(u_i, v_i) X_i)^{-1} X_i^T W(u_i, v_i) Y_i, \tag{2}$$

where $\hat{\beta}(u_i, v_i) = \begin{matrix} \hat{\beta}_0(u_i,v_i) \\ \cdots \\ \hat{\beta}_r(u_i,v_i) \end{matrix}$ - vector of parameters estimations $\hat{\beta}_r(u_i, v_i)$, $X_i = \begin{matrix} x^i_{11} & \cdots & x^i_{r1} \\ \cdots & \cdots & \cdots \\ x^i_{1n_i} & \cdots & x^i_{rn_i} \end{matrix}$ - matrix of size $n_i * r$ of values of input factors F_k at observation points for i-th area, $Y_i = \begin{matrix} y^i_1 \\ \cdots \\ y^i_{n1} \end{matrix}$ - vector of values of the dependent variable y^i at observation points, $W(u_i, v_i)$ - matrix of weights for each point (u_i, v_i). The elements of the matrix $W(u_i, v_i)$ are chosen so that observations near the point (u_i, v_i) have a greater weight than observations that are far away. Matrix has the following form:

$$W(u_i, v_i) = diag\{w_{i1}, ...w_{in_i}\}, \tag{3}$$

where w_{ij} - weight of observation at point j for point i. Weight can be calculated as follows:

$$w_{ij} = e^{-0.5(d_{ij}/h)^2}, \tag{4}$$

where d_{ij} - is Euclidean distance between points i and j, and h - a parameter that affects the bandwidth of a geographically weighted regression.
Matrix $M = X_i^T W(u_i, v_i) X_i$ of size $r * r$, undergoing the operation of reversal in (2), for certain sets of observations may become degenerate. This happens because sets of dummy variables were used for qualitative factors. Therefore, instead of the usual inversion in (2), the pseudo-inversion operation of Moore-Penrose was used:

$$M = UDV^T, M^+ = VD^+U^T, \tag{5}$$

where U and V - unitary matrices of order r, consisting of left and right singular vectors, respectively, and D is a diagonal matrix of size $r * r$, containing singular numbers on the main diagonal.

2 Methods of kriging family

To predict the value of Z_i in some point of interest i of area Ω_i using methods of kriging family, it is necessary to average existing observations (6):

$$Z_i = \sum_{k=1}^{n_i} \alpha_k(u_i, v_i) Z_k, \tag{6}$$

where $(u_i, v_i) \in \Omega_i$ - point of interest with coordinates (u_i, v_i), where the value of Z_i (speed) is predicted, Z_k- value of transport speed observation in (u_k, v_k), $\alpha_k(u_i, v_i)$- unknown weight coefficient, n_i - the number of observations Z_k in Ω_i.
To predict the values of Z_i one needs to find $\alpha_k(u_i, v_i)$. At the first stage, it is necessary to analyze the spatial correlation structure of the initial set of observations. To do this, one should use a statistical moment such as a variogram [6]. Its values are explicitly included in the kriging equations.
Variogram is a correlation measure for two values of the observed variable in points Z_k and Z_j, located at a distance $d(Z_k, Z_j) = h$ from each other:

$$\gamma(h) = 0,5Var[Z_k - Z_j] = 0,5E[Z_k - Z_j]^2. \tag{7}$$

For N_h observation points, located at a distance h from each other,
$A_h = \{(Z_k, Z_j)|d(Z_k - Z_j) = h\}$:

$$\gamma(h) = \frac{1}{2N_h} \sum_{(Z_k, Z_j) \in A_h} [Z_k - Z_j]^2. \tag{8}$$

To construct an experimental semivariogram, it is necessary to group all pairs of measurement points by distance and calculate the semivariogram values for all groups using formula (8). A permissible range of distances, called a lag, is used during variogram calculation. Pairs of points are grouped using lag value and it provides some reduction in the effect of the emissions on its values. Kriging methods require knowledge of the variogram values for all distances. For this purpose, a theoretical semivariogram model is constructed $\hat{\gamma}(h)$ - a function that approximates the values of the experimental variogram. The functions that were used to approximate the experimental variogram are given in [7]. These models are functions defined up to parameters. One of the parameters is the correlation radius, which is the limiting distance between points, at which a correlation effect is still observed. After the stage of the semivariogram calculation, one should proceed to the direct prediction of the values of the observed quantity.
It is possible to decompose observed variable $Z(u, v)$ into the sum of the deterministic and stochastic components: $Z(u, v) = m(u, v) + R(u, v)$. The methods of the kriging family differ with each other by propositions about the form of the deterministic component - mean of the function $m(u, v)$.

2.1 Universal kriging

To calculate the estimates in this paper, universal kriging was applied. Universal kriging [5] suggests that the deterministic component $m(u, v)$ of $Z(u, v)$ is a linear combination of certain basis functions $f_p(u, v)$, with coefficients $\lambda_p(u, v)$. These coefficients are assumed to be constant in Ω_i.

$$m(u, v) = \sum_{p=0}^{T} \lambda_p(u, v) f_p(u, v), \tag{9}$$

where $(u, v) \in \Omega_i$, $f_p(u, v)$ - selected basis functions, $f_0(u, v) = 1, \lambda_p(u, v)$ - unknown coefficients, $T + 1$ - the number of used basis functions.
It should be noted that kriging methods, including universal one, allow calculating local estimates of the function $Z(u, v)$, namely, to use for estimation only those measurements that are located in some neighborhood Ω_i of point (u, v).
The solution of the problem of finding unknown weights is carried out using the minimization of the functional $\sigma_L^2(u, v)$, where in addition to the variation of the estimation error by the kriging method, the unbiasedness of the estimate with weight coefficients μ_p are included (the Lagrange multipliers):

$$\begin{aligned}\sigma_L^2(u_i, v_i) = \sum_{k=1}^{n_i}\sum_{j=1}^{n_i}\hat{\alpha}_k(u_i, v_i)\hat{\alpha}_j(u_i, v_i)(\sigma_Z^2 - \hat{\gamma}_{kj}) - \\ -2\sum_{k=1}^{n}\hat{\alpha}_k(u_i, v_i)(\sigma_Z^2 - \hat{\gamma}_{ik}) + \sigma_Z^2 + \\ +2\mu_0(u_i, v_i)(1 - \sum_{k=1}^{n_i}\hat{\alpha}_k(u_i, v_i)) + \\ +2\sum_{p=1}^{T}\mu_p(u_i, v_i)(f_p(u_i, v_i) - \sum_{k=1}^{n_i}\hat{\alpha}_k(u_i, v_i)f_p(u_k, v_k))\end{aligned} \tag{10}$$

where $\hat{\gamma}_{kj} = \hat{\gamma}(d(Z_k - Z_j))$.
The system of universal kriging equations, obtained by differentiating the variation (10) by weights $\hat{\alpha}_i$ and coefficients μ_p and equating the derivative to zero, looks as follows:

$$\begin{cases}\sum_{j=1}^{n_i}\hat{\alpha}_j(u_i, v_i)\hat{\gamma}_{kj} + \mu_0(x) + \sum_{p=1}^{T}\mu_p(u_i, v_i)f_p(u_k, v_k) = \hat{\gamma}_{ik}, k = 1, ..., n_i, \\ \sum_{k=1}^{n_i}\hat{\alpha}_k(u_i, v_i) = 1, \\ \sum_{k=1}^{n_i}\hat{\alpha}_k(u_i, v_i)f_p(u_k, v_k) = f_p(u_i, v_i), p = 1, ..., T.\end{cases} \tag{11}$$

The variation of universal kriging can be calculated from formula (10) using the first part of system (11) as follows:

$$\sigma^2 = \sum_{k=1}^{n_i}\hat{\alpha}_k(u_i, v_i)\hat{\gamma}_{ik} + \mu_0(u_i, v_i) + \sum_{p=1}^{T}\mu_p(u_i, v_i)f_p. \tag{12}$$

2.2 Kriging with external drift

In many practical problems of spatial estimation, the value of the interpolated variable may be accompanied by additional information, presented in the form of external parameters, set over the entire field of observation. Under certain conditions, additional information may contribute to improving the quality of interpolation of the values of the considered value. The main condition for the ossibility and expediency of using additional information is its correlation with the main variable being assessed.

Kriging with external drift can be considered as a modification of universal kriging - if in the case of predictor variables only measurement coordinates are used. However, in the general case, kriging with external drift assumes that the deterministic part of the measurement can be defined as a combination of functions of some auxiliary variables $g_p(u, v)$, known up to coefficient values $\lambda_p(u, v)$:

$$m^{KED}(u, v) = \sum_{p=0}^{T} \lambda_p(u, v) g_p(u, v), \tag{13}$$

where $(u, v) \in \Omega_i$, $g_p(u, v)$ - known functions of auxiliary variables, $g_0(u, v) = 1, \lambda_p(u, v)$ - unknown coefficients, $T + 1$ number of auxiliary variables.
By constructing the corresponding Lagrangian (by analogy with the functional of the universal kriging method (10), its differentiation over all unknown variables and equating the corresponding derivatives to zero, a system of kriging equations with external drift (14) is obtained:

$$\begin{cases} \sum_{j=1}^{n_i} \hat{\alpha}_j(u_i, v_i)\hat{\gamma}_{kj} + \mu_0(x) + \sum_{p=1}^{T} \mu_p(u_i, v_i) g_p(u_k, v_k) = \hat{\gamma}_{kj}, k = 1, ..., n_i, \\ \sum_{j=1}^{n_i} \hat{\alpha}_j(u_i, v_i) = 1, \\ \sum_{j=1}^{n_i} \hat{\alpha}_j(u_i, v_i) g_p(u_k, v_k) = g_p(u_i, v_i), p = 1, ..., L. \end{cases} \tag{14}$$

Most often, in the kriging method with external drift, the variables that have a close to linear dependence with the estimated value and vary quite smoothly on the studied area are taken into account as additional.
In this paper, the same variables are used as additional variables for kriging with external drift as for the method of geographically weighted regression In this paper, the same variables are used as additional variables for kriging with external drift as for the geographically weighted regression method, which are mentioned in (1).

3 Processing of GPS tracks of road users

The efficiency of the above-described GWR model was verified based on statistical information on the movement of the Novosibirsk city private transport. The initial data was an array containing more than one million records of the instantaneous speeds of users of one of the known GPS navigators. Each element of the data array contained:

- user coordinates,
- unique user identifier,
- direction of movement (azimuth) of the user,
- instantaneous speed,
- time of fixing the speed.

The topology of the transport network of the Novosibirsk city was obtained from the open source OpenStreetMap [3]. It contains information on roads, intersections and their attributes, such as the number of lanes, the maximum allowed speed, road type, max. throughput, type of intersection regulation and allowed turns.
At the first stage, the analysis and processing of the initial data was carried out. All observations of road users were processed and grouped by time intervals (1 hour) on weekdays and workdays. At the same time, the vehicles were linked to the roads, taking into account the direction of the vehicle, the type of road, the number of lanes and the direction of the road. The data received from cars standing at the curb was not taken into account.
To apply the above methods, it is necessary to select the points of interest (u_i, v_i). In addition, it is necessary to determine the observation sets that will be used to construct the regression at each of the points (u_i, v_i). Thereafter, the second stage of data preparation consisted of the formation of a list of points of interest (u_i, v_i) and subsamples of the initial sample of observations obtained at the previous stage for each such point.
The topology of the transport network was taken into account. In all the considered methods, the observations included observations with the point of interest on the same road, taking into account the direction of interest. However, for methods of geographically weighted regression and kriging with external drift, observations were also added to the sample, which are located on roads adjacent to the road point of interest (having common intersections).
A study and comparison of the presented grouping methods in terms of the quality of regression models and computational costs was carried out. The following quality criteria was selected: the adjusted determination coefficient , the residual sum of squares, the maximum and minimum values of the residuals.

4 Investigation and comparison of models

To illustrate the results of the study of transport velocity models, the most indicative time interval (from 18 to 19 hours on weekdays) was selected. This period of time is interesting because the transport network is experiencing peak loads due to the massive movement of city residents from work to their homes. As points of interest (u_i, v_i), in which velocities were predicted, random 2000 observations on the most popular roads were taken from the available initial sample. Whole initial data contained 692,634 observations.

In cases of geographically weighted regressions, for each local area, input factors that have the same values for the entire set of observations in the corresponding subsample were excluded. Similarly, factors were taken into account in the interpolation of values using the kriging method with external drift. In the case of the universal kriging method, factors were clearly not taken into account in the model.
It was necessary to choose the optimal form of weight function and values of it's parameters to calculate the weights $W(u_i, v_i)$. This parameters affects the weight

of the observations and should be chosen taking into account the density of observations near the local points under study. The dependence of the adjusted coefficient of determination of the geographically weighted regression model R^2 on the parameter h for different weight functions was studied. The study showed that the best model corresponded to weight function (4) and $h = 70$.
In the universal kriging method, the model that best approximates the experimental variogram was chosen as an exponential model [7], a distance of 205 meters was chosen as the optimal correlation radius, and the type of trend model was linear. For the method kriging with external drift was also carried out to investigate the dependence of quality vehicles velocity estimates obtained considering external factors - the characteristics of the roadway on the magnitude of the correlation radius exponential model variogram results are shown in Table 1.

Table 1: Comparison of results of speeds evaluation by kriging with external drift at different correlation radiuses (N = 2000)

Correlation radius	**105 meters**	**205 meters**	**425 meters**
Adjusted determination coefficient	0.36	0.53	0.42
Residual sum of squares	840968	589053	783514
Minimum value of residuals	-81.40	-79.53	-81.70
Maximum value of residuals	75.44	63.51	75.06

The closest to the actual speed values were obtained with a correlation radius of 205 meters (similar to universal kriging). Reducing the correlation radius to 105 meters leads to a significant decrease in the number of points in the samples, which negatively affects the quality of the estimates. However, an increase in the correlation radius to 425 meters leads to the use of a larger number of points and, consequently, to their noticeable influence on the estimates, although in practice there may be no effect on the speed at such a significant distance. Table 2 shows the results of comparing the methods of estimating speeds with the optimal parameters of the models.

Table 2: Comparison of methods for estimating speeds (N=2000)

Method	**GWR**	**Universal kriging**	**Kriging with external dreft**
Adjusted determination coefficient	0.75	0.68	0.53
Residual sum of squares	356613	454452	589053
Minimum value of residuals	-66.22	-66.67	-79.53
Maximum value of residuals	74.15	68.99	63.51

Figures 1, 2, 3 show the results of estimating the speed of the traffic flow on the streets of the Novosibirsk city using the methods presented in this paper.

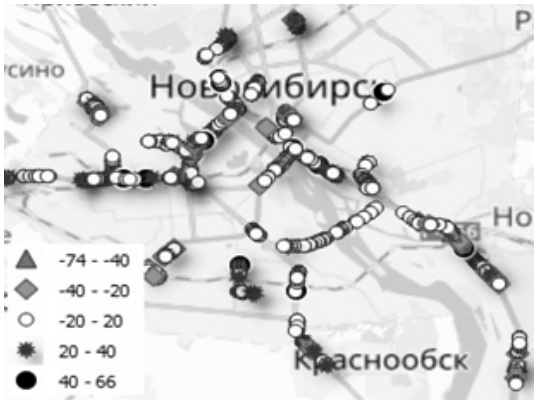


Figure 1: Residuals between actual and calculated speeds using geographically weighted regression

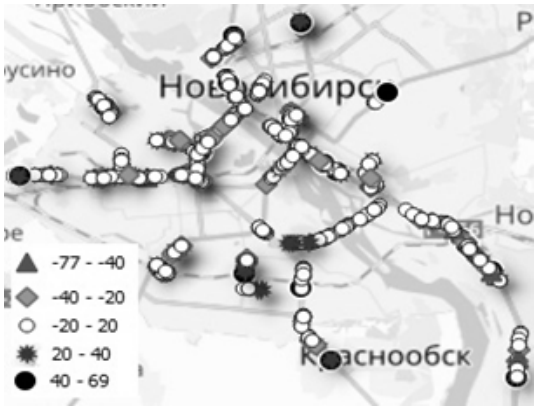


Figure 2: Residuals between actual and calculated speeds using the universal kriging

All models well estimate the speed of traffic in the Novosibirsk city. The most accurate results are provided by a model based on geographically weighted regression (Fig. 1). The universal kriging model (Fig. 2) demonstrates a somewhat less high correlation between estimated and actual values, however, this model is independent of factors and uses only the values of the observed variable. The model obtained by estimating the values of kriging with external drift (Fig. 3) demonstrates the least high correlation between the actual and estimated values, despite the fact that this model, like the model based on the geographically weighted regression method, takes into account the values of external factors - characteristics of the roadway. This is probably due to the fact that the relationship between speed and some factors is not linear, or the values of the factors change unevenly on the study area. However, it is worth noting that the magnitude of the largest deviations of the estimated values from the actual values in the kriging family methods is less than in the geographically weighted regression method. The external drift kriging method is less prone to significant outliers.

Figure 3: Residuals between actual and calculated speeds using kriging with external drift

Conclusion

It was proposed to use the geographically weighted regression method and methods of Kriging family for constructing the model of private transport speeds in real urban conditions in the paper. The quality of the estimation of the universal kriging method from the types of the trend and variogram models, as well as the quality of the estimation of the geographically-weighted regression method from the form of the weighing function, were investigated. A study of models on real data of Novosibirsk was carried out. During the study, the optimal bandwidth for GWR and lag for Kriging model were obtained. This study showed the expediency of using the proposed methods for estimating the speeds of transport.

References

[1] Timofeev V.S., Teselkina K.V., Veselova A.S. Development and Research of Transport Speed Models Using the Methods of Geo-statistical Data Analysis, Novosibirsk, Russia // 14th International Scientific-Technical Conference on Actual Problems of Electronic Instrument Engineering, APEIE - 2018, pp. 315-319.

[2] Geographically Weighted Regression: The Analysis of Spatially Varying Relationships, A. Stewart Fotheringham, Chris Brunsdon, Martin Charlton, 2002

[3] https://www.openstreetmap.org

[4] Searle S. Linear Models, New York: Wiley, 1971

[5] Demianov V. V. Geostatistika teoriya i praktika : [monograph] / V. V. Demianov. E. A. Savelyeva; In-t problem bezopasnogo razvitiya atomnoy energetiki Ran - M.: Nauka. 2010. - p. 327

[6] Kovalevskij E.V. Geologicheskoe modelirovanie na osnove geostatistiki // EAGE - 2011. p.117

[7] Armstrong M. Basic Linear Geostatistics, 1998

Study of the properties of geometric ABOD-approach modifications for outlier detection by statistical simulation

Nikita S. Oleinik and Vladislav Yu. Shchekoldin
Novosibirsk State Technical University, Novosibirsk, Russia
e-mail: olenik.2015@stud.nstu.ru, raix@mail.ru

Abstract

In the article the problem of outlier detection in multidimensional data is discussed. The capabilities of ABOD-approach as an element of a family of spatial-based methods are being studied. For improvement of outlier detection quality in multidimensional data the modification of the criterion based on cumulative curve analysis is proposed. The comparison of results obtained by classical and modified algorithms is provided.

Keywords: multivariate statistical analysis, outliers, ABOD, statistic simulation.

Introduction

In contemporary studies of complex technical and economic processes one has often to deal with big data formed by the values of numerous indicators. For such multidimensional data it is necessary to develop specialized methods for collecting, processing and analyzing them, since classical approaches of statistics and econometrics could give considerably distorted results. In this case one of the problems most influenced by dimensionality is an identification of specific (anomalous) elements in the sample called outliers.

An outlier, according to D. Hawkins, is «an object that deviates from others so much as to be suspected that a different mechanism generated it» [2]. In the case of multidimensional data analysis the outlier detection will often face with so-called «curse of dimension» – the property of multidimensional spaces, meaning that as the dimension increases, there is an exponential increase in the volume of necessary experimental data, which complicates both the analytical and the algorithmic apparatus used to solve such problem [1, 6].

In the past few decades in multidimensional data analysis some algorithms based on the geometry of the mutual arrangement of observations in multidimensional space have become widespread. In previous studies (including authors [4]) it has been shown that in comparison with other frequently used algorithms ABOD-method is invariant to changes in the dimension of space, and this is a reason to intensive studying and development it in applications.

1 Problem statement

Let there be a database $D \in R^d$, where R^d is a real space, d - spatial dimension, $n = card(D)$ - database's size. We will assume that the norm $\mu : R^d \to R_0^+$ is associated with R^d. It induces the scalar product $(.,.) : R^d \times R^d \to R$. In the current study we will consider Euclidean $\|x\|_2 = \sqrt{\sum_i |x_i|^2}$ and Manhattan $\|x\|_1 = \sum_i |x_i|$ as a norm μ, because they are most frequently used when investigating the issues of detecting anomalous observations.

Since the initial set D is divided by the algorithm into subsets D_1 (regular observations), D_2 (outliers) and, possibly, D_3 (intermediate), the main hypothesis $H_0 : x \in D_1$ means that the observation x is regular. It can be rejected in favor of the alternative $H_1 : x \in D_2$ or, maybe $H_2 : x \in D_3$

The purpose of the study is to build a decision rule $G : D \to W$, where $W = D_1 \cup D_2 \cup D_3$ to classify sample observations into two (or, depending on the method's capabilities, into three) categories – regular observations, outliers, and possibly intermediate observations with unclear situation arising, which can be clarified by applying more accurate and/or statistically more robust methods.

2 Theoretical basis

The family of geometric methods is developing nowadays and becoming more and more popular, so, according to [3], most popular among them are DiBOD (Distance Based Outlier Detection), DeBOD (Depth Based Outlier Detection), ABOD (Angle Based Outlier Detection), and others.

The principle of DiBOD-method is as follows [1]: it is necessary for each observation to find the number of other observations lie in some its neighbourhood, and then observations for which the number of neighbours is significantly less than the average will be considered as outliers. The benefits include ease of implementation of the algorithm and a small number of computational operations. But it has a serious drawback: it is weakly invariant to multidimensionality.

Method DeBOD based on the concept of «depth of space», suggested by P. Rousseeuw [8]. According to it outliers will be observations that correspond to the smallest depth value. This approach is applicable when one need to process a large amount of data in a short time, but also, like DiBOD, is weakly invariant to multidimensional problems.

Unlike DiBOD and DeBOD, ABOD-method is invariant to multidimensionality, and the method itself is based on estimation the multidimensional angles at which all other observations would be visible from a certain point in space corresponding to a particular observation in the sample.

In this case on the basis of the following expression the estimation of the angle variance for observation x is carried out [3]:

$$ABOF(x) = \underset{y,z \in D}{VAR} \left(\frac{(\overline{xy}, \overline{xz})}{||\overline{xy}||^2 * ||\overline{xz}||^2} \right), \quad (1)$$

where ABOF(x) is a function which estimates the degree of anomality for observation x, VAR(.) – variance function, y, z – points in multidimensional space, selected from database D, (. , .) – scalar product of two vectors represented by two points from the database, ‖.‖ – norm (length) of the corresponding vector in multidimensional space.

Observations whose maximum values of type (1) are minimal in the sample are potentially considered outliers.

In order to determine that the value in (1) is small enough to consider the corresponding observation as an outlier it is necessary to apply certain statistical test. From a wide variety of such tests we will focus on the following:

- Chauvenet test;
- maximum relative deviation (MRD) test;
- test based on cumulative curve analysis (CCA-test).

Chauvenet and MRD-tests are based on the assumption of data's normal distribution. This assumption in the analysis of actual data is appeared infrequently, primarily because the data homogeneity is not hold. However it is important to understand that even in such situation with sufficiently large sample sizes the methods developed under the assumption of normality could bring consistent results. In this case it is said that normality shows asymptotically.

When using Chauvenet test the suspected value of the sample x^* delivers maximum to the absolute deviation of the mean [4], namely

$$x^* = Arg\,\max_k |x_k - \overline{x}|.$$

In this case the statistic for testing the hypothesis H_0 is $\widehat{\tau} = \frac{|x^* - \overline{x}|}{S}$ ($\overline{x}$ - arithmetical mean, S- standard deviation), and if inequality $\widehat{\tau} > |u_{\frac{1}{4n}}|$ is hold, when the null hypothesis is rejected, and observation x^* is recognized as an outlier.

When using MRD-approach the test statistic remains the same as in Chauvenet test [4], but the decisive rule is changing as follows:

$$\tau_{\alpha,n} = \frac{t_{1-\alpha,n-2}\sqrt{n-1}}{\sqrt{n-2+t^2_{1-\alpha,n-2}}}.$$

where $\alpha - quantile\ of\ Student's\ distribution$, $t - student\ distribution\ quantile$.

Here if $\widehat{\tau} < \tau_{5\%,n}$ then suspected observation considered as regular, else if $\widehat{\tau} > \tau_{0.1\%,n}$ it becomes an outlier, and otherwise it occupies an intermediate position, and will be an outlier if one has some additional considerations on it.

In order to provide the invariancy relative to the distribution of analyzed data and to take into account their internal structure, not relying on the critical values associated with certain distributions (like normal distribution in Chauvenet and MRD tests), one could apply the test based on the cumulative curves analysis. To construct a specific cumulative curve it is necessary to determine its specification and after that to estimate its parameters by using one of various approaches, for example, Gauss-Markov least squares [5].

In the current study the cumulative curve specification was chosen as Ballou-Pareto type. Within this class of functions the most successful appeared to be the parameterization of Ballou-Pareto type III ($f(x) = \frac{(\alpha+1)x^\gamma}{\alpha+x^\gamma}$), since it often provides the minimum value of the residual sum of squares for the problem being solved [5, 7, 9].

The next step of analysis is the implementation of the integral method [7, 9] according to which the cumulative curve is going to be divided into three parts corresponding to outliers, intermediate observations and regulars.

There are various standard outlier detection software, such as Statistica, SPSS, etc. One of the main problems of them is that they realize only some classical methods based on normal distribution, which mostly result poorly, or even lead to contradictory conclusions for multidimensional tasks. Besides in academic versions applying in the educational process they do not allow analyzing big data. Therefore it was decided to develop the original software package, which includes

- module to input and generate data for analysis;
- module to apply the classical outlier detection tests;
- module to apply some modifications of classical methods (ABOD and CCA);
- statistical simulation module;
- module for analytical report construction (tabular and graphical presentation of results).

In this way the developed approach and its modifications will have both analytical and algorithmic support which will allow studying their properties and establishing the scope of applicability.

3 Experimental results

In order to ensure statistical correctness of the results obtained it is necessary to explore the proposed algorithm on model problems by simulation.

For studying purposes let us model the samples with data dimensions as $d = 3$, 4, and 5, and sample size equal to $n = 100$. The initial data are going to be simulated uniformly on the basis of the corresponding equations for multidimensional spherical coordinates. For conventionally regular observations the polar radius ρ lies within interval from 0 to 2, for outliers – from 3 to 4 (which means from 150% to 200% out of initial values, which should provide their significant difference from regular observations). For each dimension the share of outliers will vary from 5% (standard statistical error) till 15%. The results to comparison of the considered methods' efficiencies are the shares (in percentage) of correctly detected outliers which are averaged for all replications (number of replications is $R = 1000$).

When comparing the results obtained by classical methods and with the help of the algorithm proposed in the paper for estimating the mutual arrangement of

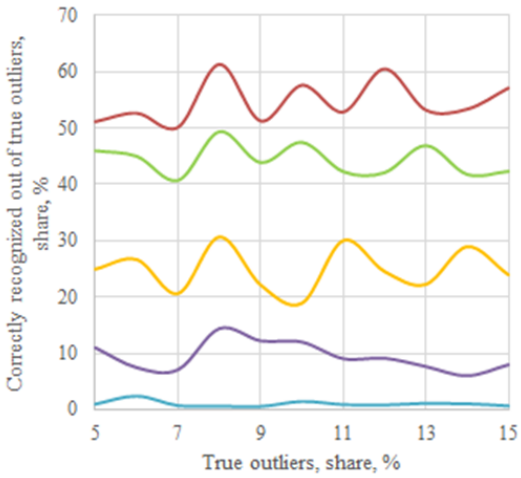


Figure 1. Comparison of shares of correctly identified outliers in 3D-space

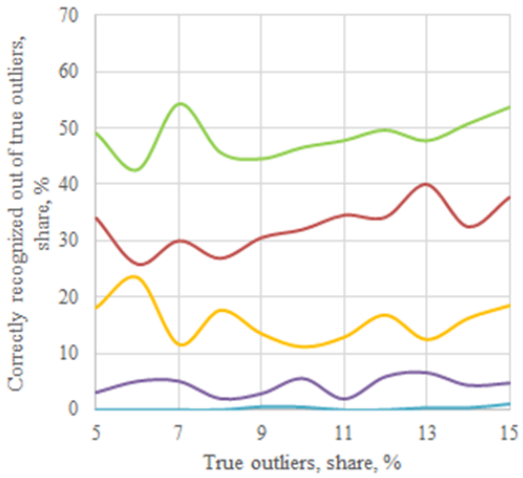


Figure 2. Comparison of shares of correctly identified outliers in 4D-space

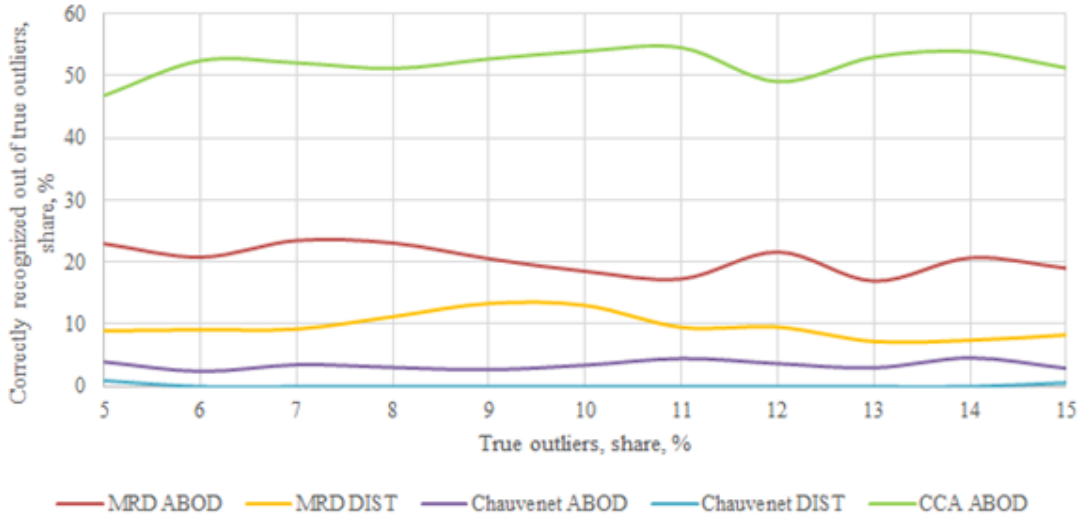


Figure 3. Comparison of shares of correctly identified outliers 5D-space

observations in d-dimensional space two options were considered: on the basis of sum of distances to all observations (denoted as DIST-modification) and on the basis of angular variance (ABOD-modification).

In Fig. 1-3 comparisons of the true irregular observations in the sample and outliers, identified during the simulation using Chauvenet, MRD and CCA-tests, used with DIST- and ABOD- modifications appropriately are depicted. From Fig. 1 it can

be seen that the proportion of correctly detected outliers using ABOD-modification is higher than using the DIST. For example, for MRD-test this difference ranges from 1,75 to 3,02 times, and for Chauvenet from 1,58 to 4,76 times, i.e. it grows with the increasing share of simulated outliers. At the same time it can be found that the more technically complicated CCA-method does not work well (its results are worse than for MRD-ABOD by 23% on average). It seems reasonable since the effectiveness of the CCA-method should become noticeable at high dimensions of the analyzed data space.

From Fig. 2 it is clear that there is also a tendency that the share of detected outliers using ABOD-modification is higher than using DIST. For example for MRD this difference ranges from 1,10 to 3,25 times, and for Chauvenet from 1,88 to 9,33. However, unlike the previous case, CCA-method provides a share of correctly recognized outliers higher than MRD from 1,19 to 1,81 times. It comes out due to the fact that with the increase in the space dimension it becomes more difficult to recognize outliers for MRD and Chauvenet tests, which confirms their weak invariance to changes in the dimension of the observation space.

In fig. 3 that when increasing the dimension, the share of detected outliers by the CCA algorithm remained almost unchanged (no more than +23% when comparing five-dimensional and four-dimensional spaces, or from + 2% to + 29% for five-dimensional and three-dimensional, respectively) compared to other algorithms that once again confirms its dimensionality invariance.

Analyzing the results one could see that when comparing four- and three-dimensional spaces MRD-algorithm with ABOD-modification (MRD-ABOD) on average finds 21% less outliers, MRD with DIST- modification (MRD-DIST) - 9,2% less , Chauvenet-ABOD - 5,3% less, Chauvenet-DIST is 0,8% less and only CCA reveals 3,9% more on average.

When comparing the results for five- and four-dimensional spaces, it can be noted that MRD-ABOD finds 12% less outliers (averaged on true outliers share), MRD-DIST - on 5,7% less, Chauvenet-ABOD - on 0,7% less, Chauvenet-DIST on 0,1% less, and only CCA reveals on 3,6% more on average.

Such results give us assurance that suggested ABOD-modification appears to be more efficient than usually used DIST-modification. On the other hand the application of CCA-approach make it possible to reveal bigger share of true outliers than other methods because it could adapt better to the inner structure of multispaced data.

Conclusions

In the study of ABOD-modification it was found that on the simulated data it shows more accurate results, since on identical samples it provides the error of the first type significantly less than for the method based on distances between observations. It is also worth noting that as the dimension increases the results obtained on the basis of CCA-approach remain at the same level, unlike other algorithms when the error of the first kind of which increases drastically.

Based on the analysis performed it should be concluded that the combination of ABOD-method and CCA-algorithm is promising for solving classification problems including in identifying anomalous observations. Therefore, their further study and development seems necessary and challenging.

References

[1] Knorr E., Ng R. (1998). Algorithms for Mining Distance-Based Outliers in Large Datasets. Proceedings of the 24rd International Conference on Very Large Data Bases. P. 392-403.

[2] Hawkins D. (1980). Identification of Outliers. New York, Chapman and Hall.

[3] Kriegel H., Schubert M., Zimek A. (2008). Angle-Based Outlier Detection in High-dimensional Data. Proc. of the 14th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD'08). Las Vegas, NV. P. 444-452.

[4] Lin L., Sherman P. (2007). Cleaning data the Chauvenet way. Proc. SouthEast SAS Users Group. Paper SA11. P. 1-11.

[5] Lyssenko M., Shchekoldin V. (2018). Development of classification methods based on cumulative curves analysis. 14th International Scientific-Technical Conference on Actual Problems of Electronic Instrument Engineering (APEIE-2018). T. 1, Vol. 4. P. 164-167.

[6] Novak E., Ritter K. (1997). The curse of dimension and a universal method for numerical integration, multivariate approximation and splines. Internat. Ser. Numer. Math., Birkhäuser, Basel. Vol. 125. P. 177-187.

[7] Oleinik N., Shchekoldin V. (2018). Identification of abnormal observations in the large dimension data based on the geometric ABOD-approach. Proceedings of Science, Technologies, and Innovations, Novosibirsk. Vol. 2. P. 253-257. (in Russian)

[8] Rousseeuw P., Aelst S., Driessen K., Agullo J. (2004). Robust multivariate regression. Technometrics. Vol. 46. P. 293-305.

[9] Shchekoldin V. (2016). Developing the risk classification based on ABC-analysis of possible damage and its probability. Proceedings of IFOST-2016, Novosibirsk. T. 1, Vol. 1. P. 317-319.

On the clustering task of Big Data in medicine and neurophysiology

Yurii Mezentsev[1], Olga Razumnikova[1], Irina Tarasova [2], Olga Trubnikova [2]

[1] *Novosibirsk State Technical University, Novosibirsk, Russia*

[2] *Research Institute for Complex Issues of Cardiovascular Diseases, Kemerovo, Russia*

e-mail: mesyan@yandex.ru, razoum@mail.ru, iriz78@mail.ru, olgalet17@mail.ru

Abstract

The NP-hard clustering problem as applied to the data of neurophysiological studies (indicators of postoperative cognitive dysfunction) is considered. Variants of the problem of clustering in the form of mixed integer programming, including the use of continuous relaxation, reducing the complexity of solutions without loss of accuracy are given. The results of computational experiments on real data using the software implementation of the algorithm of binary cuts and branchings are presented. They demonstrate the high efficiency of the developed toolkit.

Keywords: clustering, minimax quality criterion, additive criterion, linear relaxation, binary cut and branch algorithm, detection of postoperative cognitive dysfunction.

Introduction

There are many works on applied statistics and methods of discrete optimization, which are devoted to the development of various modifications of applied problems of clustering [1, 2, 3]. Most of them belong to the category of intractable problems of mixed integer programming [2]. This study is aimed to solving the applied clustering problem by identifying and comparing the substantive patterns of partitions of various sets of objects into subsets due to comparing such partitions by EEG characteristics in patients before coronary bypass surgery and in the postoperative period.

1 Statements of the Clustering Problem

This article is presented an approach based on the use of tools (models and methods) of mixed integer linear programming (milp). A similar problem has been formulated early [4, 5]. The directions of connections between objects within each cluster are not fixed and must be determined. Cluster intersections by objects (and objects by clusters) are not allowed. The most rational and most often used measure is the sum of the distances between all pairs of objects within a cluster [1]. We present a formal formulation of the optimal clustering problem for this case.

We introduce the notation: $i,j = \overline{1,n}$,- object numbers, $l,k = \overline{1,m}$. - cluster numbers. $c_{i,j}^k$ - distances between objects i and j in cluster k. Then the task of clustering is to determine the Boolean variables $x_{i,j}^k$. Define the variables y_i^k (identifying the ownership of objects $i,j,i,j = \overline{1,n}$, into cluster $k,k = \overline{1,m}$) under certain conditions:

$$y_i^k = \begin{cases} 1 \text{ if the object } i \text{ belongs to a cluster } k \text{ ,} \\ 0 \text{ otherwise, } i = \overline{1,n}, k = \overline{1,m}, \end{cases} \tag{1}$$

$$\sum_{k=1}^{m} y_i^k = 1,\ i = \overline{1,n}, \tag{2}$$

We also define dependent variables $x_{i,j}^k = y_i^k \cdot y_j^k$ and linearizing inequalities [8]: $0 \leq y_i^k + y_j^k - 2x_{i,j}^k \leq 1$, $k = \overline{1,m}$, $i,j = \overline{1,n}$, $i \neq j$, which by means of an asymmetric distance matrix, are converted into:

$$0 \leq y_i^k + y_j^k - x_{i,j}^k - x_{j,i}^k \leq 1, k = \overline{1,m},\ i,j = \overline{1,n},\ i \neq j. \tag{3}$$

$$x_{i,j}^k = \begin{cases} 1 \text{ if the objects } i,j \text{ belongs to a cluster } k\ :\ y_i^k = 1, y_j^k = 1, \\ 0 \text{ otherwise, } i,j = \overline{1,n},\ i \neq j, k = \overline{1,m}. \end{cases} \tag{4}$$

By adding conditions to the task that implement the minimax criterion:

$$\sum_{j=1}^{n} \sum_{i=1}^{n} c_{i,j} x_{i,j}^k \leqslant \lambda,\ k = \overline{1,m}, i \neq j, \lambda \to \min. \tag{5}$$

we obtain the desired formalization (1) - (5) of the task named in the section title.

In addition to criterion (5), depending on the meaning of the clustering problem, in some cases the additive criterion is more acceptable:

$$\sum_{j=1}^{n} \sum_{i=1}^{n} c_{i,j} x_{i,j}^k = \lambda^k,\ k = \overline{1,m},\ i \neq j, \sum_{k=1}^{m} \lambda^k \to \min \tag{6}$$

where λ^k - sum of distances between all pairs of objects in a cluster $k,k = \overline{1,m}$.

Solving problem (1) - (4), (6) allows finding partitions of a set of objects with given distances between all pairs of objects into a given number (m) of subsets (clusters), which guarantees minimization of the sum due to the minimum of total distances between all pairs of objects across all clusters.

2 Assessments of the Computational Complexity of the Clustering Problem and Relaxation Possibilities

Note that even with a significant simplification of the limiting conditions of the above problem, the NP-hard problem of mixed integer programming is obtained (see, for

example [6, 7]). Let us show how to somewhat ease the complexity of tasks (1) - (5) and (1) - (4), (6). For this we use relaxation on auxiliary Boolean variables $x_{i,j}^k$ removing the integer conditions (4). Instead, we introduce the boundaries of changes in continuous variables:

$$0 \leqslant x_{i,j}^k \leqslant 1, k = \overline{1,m},\ i,j = \overline{1,n},\ i \neq j.] \tag{7}$$

Relaxation of tasks (1) - (5) and (1) - (4), (6) we denote as (1) - (3), (5), (7) and (1) - (3), (6) - (7). We note a decrease in the number of Boolean variables in relaxed tasks by $m \cdot n^2$. Thus, the total number of Boolean variables (1) in problems (1) - (3), (5), (7) and (1) - (3), (6) - (7) is equal to $m \cdot n$ in the presence of $m \cdot n^2 + 1$ continuous variables (7) and (5) versus $m \cdot n^2 + m \cdot n$ Boolean variables in (1)-(5) and (1-4), (6) tasks.

The difference is very significant when applied in practical applications of the tasks presented. The success of the practical application of conditionally exponential algorithms strongly depends on the actual number of integer variables. So, relaxation (1) - (3), (5), (7) and (1) - (3), (6) - (7) have significant advantages over the settings (1) - (5) and (1)) - (4), (6). At the same time, continuous variables, regardless of their number in any milp, only a few make them "heavie". For example, with 3-clustering a group of 40 patients (splitting a group into 3 clusters) for the relaxed task (1) - (3), (5), (7) we have $m \cdot n = 3 \cdot 40 = 120$ Boolean variables, in the original problem (1) - (5) - $m \cdot n^2 + m \cdot n = 3 \cdot 40^2 + 120 = 4920$ Boolean variables.

3 Algorithms of Solutions

For presented above options of the NP-difficult tasks clustering, there are no theoretically efficient algorithms. However, for practical applications, algorithms are sufficiently developed, which can be referred to as conditionally exponential. An example of such an algorithm may be the binary cut and branch algorithm [9, 10]. Its software implementation was used to find solutions to the above clustering problems (1)-(3),(5),(7) and (1)-(3),(6)-(7).

4 Application of the Developed Clustering Instrumentation to Detect Postoperative Cognitive Dysfunction

A series of computational experiments on clustering EEG data based on the above statements (1) - (3), (5), (7) and (1) - (3), (6) - (7) and the corresponding discrete optimization algorithms [9, 10] identified the high efficiency of this tool. The tasks (1) - (3), (5), (7) using the minimax criterion and the adapted algorithm were used to analyze indicators of the neurophysiological status of patients who underwent direct myocardial revascularization under conditions of cardiopulmonary bypass. Neurophysiological examination was performed 3-5 days before the operation and 7-10

days after coronary artery bypass surgery. A detailed description of the sample of patients and the assessment of EEG and cognitive status presented earlier [11, 12].

For clustering into three groups, the total values of resting EEG power with closed eyes in the theta1 (4-6 Hz), alpha1 (8-10 Hz) and beta2 (20-30 Hz) bands were used. These frequency bands were considered in connection with their functional significance for attention and memory processes [13, 14] and the previously identified informational role for predicting cognitive decline in the postoperative period of coronary artery bypass surgery [12]. Taking into account the long computation time, a sample of 40 male patients (56.7 ± 5.08 years) with EEG registration before and after the operation. Also initial preoperative cognitive status indicators were taken for analysis: the sum of points on the short mental status scale (SMSS) and a complex indicator of cognitive status (CCS) as a summary characteristic of the functions of attention, and memory [15].

The parameters of the three clusters, obtained during the realization of the developed algorithm for the theta1, alpha1 and beta2 spectral power before and after the operation, are presented in Table 1.

Based on the results presented in Table 1, the variability of the cluster composition was 33.5 percents; 22.5 percents and 52.5 percents, respectively, for theta1, alpha1 and beta2 bands. At the same time, the patients set from the cluster 2 moved to the 3rd according to the classification of postoperative power of alpha1 oscillations.

Table 1: Clusters based on the minimax classification criterion for theta1, alpha1 and beta2 rhythm power indices before and after coronary artery bypass surgery

	Before surgery		After surgery			
Cluster	Metric Score	n	Cluster	Metric Score	n	Nv
Theta1 range						
1	45.58	15	1	56.18	17(10)	$(3)2cl, (4)3cl$
2	46.07	14	2	41.04	13(10)	$(3)1cl$
3	35.51	11	3	43.21	17(10)	$(2)1cl, (1)2cl$
Alpha1 range						
1	77.05	13	1	80.88	14(11)	$(1)2cl, (2)3cl$
2	83.89	14	2	78.90	14(4)	$(10)3cl$
3	74.46	13	3	73.86	12(1)	$(1)1cl, (10)2cl$
Beta2 range						
1	47.80	13	1	48.27	14(7)	$(1)2cl, (6)3cl$
2	48.99	12	2	46.60	12(7)	$(5)3cl$
3	48.22	15	3	48.12	14(5)	$(6)1cl, (1)2cl$

Note: n - the number of patients in the cluster; in parentheses show the number of patients remaining in the same cluster after the operation or moved from other clusters (cl), respectively (Nv).

To clarify the functional significance of the selected clusters, one-way Analysis of Variances (ANOVAs) was performed on each variable (i.e. spectral data in the corresponding frequency bands, SMSS, and CCS values in the group of participants) using CLUSTER (3) as a between-subjects factor. Significant effects of all analyses are shown in Table 2. Post-hoc analysis of the revealed effects was conducted using the Bonferroni procedure to adjust for multiple repeated measurements.

Table 2: ANOVA results on theta1, alpha1 and beta2 rhythms, age, and CCS variables when comparing three selected clusters

Variable	F(2,37)	P	Cluster		
			1	2	3
Theta1					
Power b/s	63.477	<0.0001	0.248	0.024	0.505
Age	5.132	0.011	58.87	57.29	53.09*
Power a/s	80.86	<0.0001	0.298	0.102	0.655
Alpha1					
Power b/s	102.34	<0.0001	1.452	0.293	0.920
CCS	2.69	0.08	0.573	0.455	0.575
Power a/s	89.79	<0.0001	0.312	0.293	0.894
Beta2					
Power b/s	54.49	<0.0001	-0.278	-0.606	-0.468
Power a/s	73.24	<0.0001	-0.261	-0.651	-0.442

Note: b/s and a/s - power scores before surgery and after surgery, correspondently; * - $p < 0.01$; the EEG power differences between clusters when $p < 0.0001$.

On the next step of analysis, the ANOVAs were performed for the STABILITY factor (2 levels, i.e., the patient belongs to the same cluster, allocated on the basis of pre- and postoperative EEG or not). Significant effect was obtained only for CCS and the beta2 rhythm: $F(1, 38) = 6.170$; $p < 0.018$ that was associated with large values of CCS in the constant cluster compared to the variable group (0.592 and 0.473). Thus, we can conclude that the developed method of clustering variables has discriminatory possibilities, since the formed clusters differ in a given criterion with a high degree of confidence. As for the functional significance of the clusters, this problem requires further study. However, the highlighted effect of the higher power of theta1 rhythm in younger patients is in line with age-related decrease in the power of low-frequency EEG rhythms [16, 17] whereas revealed lower values of CCF at the lowest alpha 1 correspond to ideas about the congruent changes of both age-related activation of the cerebral cortex at these frequencies and cognitive deficit [17, 18, 19]. The obtained relationship between the high CCS index and the STABILITY factor in the beta 2 range can be considered as a reflection of the compensatory hyperactivity

of the cortex in patients with cardiovascular diseases. Indeed, it was previously shown that a combined decrease in alpha activity and an increase in the power of beta rhythms may indicate damage to regional neural interactions [20]. There is an hypothesis that in patients with low cognitive status due to the long-existing state of chronic cerebral ischemia, the restructuring of the electrical brain activity is associated with the dominance of slow rhythms, rather than fast ones, as among those with a more permanent cognitive status [21]. However, further analysis is needed to analyze the informational value of coordinating different EEG rhythms and their local representation for the preservation of cognitive functions in patients undergoing coronary artery bypass surgery.

Conclusions

The results of computational experiments with data clustering of neurophysiological testing using the three formal statements presented in this article revealed the highest efficiency of the problem decision using the minimax criterion and the adapted for it algorithm of binary cutoff and branching. The counting time of a desktop computer to find the exact solutions to the implementation of the clustering problem for the power indices of EEG rhythms in 40 patients, with various modifications of the initial data and the criteria used, is in the range from 15 minutes to two hours. Approximate solutions can be obtained in a much shorter time, which determines good prospects for using the created instrument with a significant increase in the dimensions of realizations into tasks.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research, project no. 19-29-01017

References

[1] Ayvazyan S.A., Buchstaber V.M., Enikov I.S., Meshalkin L.D. Applied statistics: classification and reduction of dimension. M: Finance and Statistics, 1989, 608 p. (in Russian).

[2] Kelmanov A.V., Pyatkin A.V. NP-Difficulty of some Euclidean problems of partitioning a finite set of points // Journal of Computational Mathematics and Mathematical Physics. 2018. V. 58. No 5. P. 852-856. (in Russian).

[3] Kelmanov A.V., Pyatkin A.V., Khandeev V.I. On the complexity of some maximin clustering problems // Proceedings of the Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy of Sciences 2018. V. 24. No 4. P. 189-198 (in Russian).

[4] Mezentsev Y., Estraykh I. Tasks and algorithms for optimizing the schedules of parallel-sequential systems with undefined service routes // Reports of the Academy of Sciences of Higher School of the Russian Federation. Novosibirsk, NSTU Publ. 2016. No 3 (32) P.83-97. (in Russian).

[5] Mezentsev Y., Estraykh I. Problems and optimization algorithms of schedules of parallel-serial systems with undefined service routes Constructive Nonsmooth Analysis and Related Topics (Dedicated to the Memory of V.F. Demyanov), CNSA 2017 https://doi.org/10.1109/cnsa.2017.7973988

[6] Vazirani V.V. Approximation algorithms. Springer, 2001, 378 p.

[7] Pinedo M. Scheduling Theory, Algorithms, and Systems (3nd. ed.), Springer, 2008. 672 p.

[8] Avdeenko T. V., Mesentsev Y. A. Efficient approaches to scheduling for unrelated parallel machines with release dates // IFAC-Papers Online (IFAC Proceedings Volumes). 2016. 49 (12) 8 IFAC conference on manufacturing modelling, management and control MIM. P. 1743-1748.

[9] Mezentsev Y. Binary cut-and-branch method for solving mixed integer programming problems // Constructive Nonsmooth Analysis and Related Topics (Dedicated to the Memory of V.F. Demyanov), CNSA, 2017 https://doi.org/10.1109/cnsa.2017.7973989

[10] Mezentsev Y. Method of binary truncation and branching integer programming // Reports of the Academy of Sciences of Higher School of the Russian Federation. Novosibirsk, NSTU Publ. 2011. No 1(16) P. 12-25 (in Russian).

[11] Tarasova I.V., Trubnikova O.A., Kukhareva I.N., Barbarash OL., Barbarash L.S. Annual dynamics of neurophysiological parameters in patients undergoing coronary artery bypass surgery with artificial circulation // Complex problems of cardiovascular diseases. 2015. No 1. P. 18-24. (in Russian).

[12] Tarasova I.V., Trubnikova O.A., Barbarash O.L., Barbarash L.S. Electroencephalogram changes in patients with early and resistant postoperative cognitive dysfunction during coronary artery bypass surgery with artificial blood circulation // Neurological Journal. 2017. V. 22. No 3. P. 136-141. (in Russian).

[13] Klimesch W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis // Brain Research Reviews. 1999. 29. P. 169-195.

[14] Bressler S. L., Richter C. G. Interareal oscillatory synchronization in top-down neocortical processing // Curr. Opin. Neurobiol. 2015. 31. P.62-66. 10.1016/j.conb.2014.08.010

[15] Trubnikova O.A., Kagan E.S., Kupriyanova T.V., Maleva O.V., Argunova Yu.A., Kukhareva I.N. Neuropsychological status of patients with stable ischemic heart disease and factors influencing it // Complex problems of cardiovascular diseases. 2017. V.6. No1. P. 112-121. (in Russian).

[16] Cummins T. D., Finnigan S. Theta power is reduced in healthy cognitive aging // Int J Psychophysiol. 2007. 66(1) P. 10-17.

[17] Klass D. W., Brenner R. P. Electroencephalography of the elderly // J Clin Neurophysiol. 1995. 12 (2) P. 116-131

[18] Razumnikova O.M. Patterns of aging of the brain and how to activate its compensatory resources // Advances physiol. science. 2015. V.46. No2. P. 3-16. (in Russian).

[19] Belousova L.V., Razumnikova O.M., Volf N.V. Age features of communication of intelligence and EEG characteristics // J. Higher Nervous Activity. 2015. V.65. No 6. P. 699-705. (in Russian).

[20] Zappasodi F., Olejarczyk E., Marzetti L., Assenza G., Pizzella V., Tecchio F. Fractal dimension of EEG activity senses neuronal impairment in acute stroke // PLoS One. 2014. 9(6):e100199. doi: 10.1371/journal.pone.0100199.

[21] Tarasova I.V., Trubnikova O.A., Barbarash O.L., Barbarash L.S. Diagnostic value of electroencephalography indicators for early postoperative cognitive dysfunction // Creative Cardiology. 2016. V. 10. No 3. P. 220-230. (in Russian).

Problems of sub-federal budget policy in Russian Federation (The case of municipalities of the Novosibirsk Oblast)

Tatiana V. Sumskaya
Institute of Economics and Industrial Engineering of SB of RAS,
Novosibirsk State University of Economics and Management, Russia, Novosibirsk
e-mail: t.v.sumskaya@ngs.ru

Abstract

In this paper we identify the conditions of formation of the financial base of local self-government, the technique of analysis of the structure, stability of budgets and efficiency of sub-federal budget policy, calculations are carried out on materials of Novosibirsk oblast for the period 2010-2016 years . The structure of local budgets of Novosibirsk oblast is evaluated, the characteristics of heterogeneity of budget indicators before and after the transfer of funds from the regional budget are calculated. The dependence between transfers and tax and nontax revenues is analyzed; marginal effect of increasing the taxes paid to local budgets is calculated.

Keywords: local self-government, local budget, tax and non-tax revenues of the local budget, equalization of budgetary provision's differentiation, sub-federal budget policy.

Introduction

The development of intergovernmental fiscal relations in Russia in recent years goes in the direction of strengthening the formalization of the process of distribution of federal financial assistance. It also seeks to eliminate the asymmetry in the fiscal status of the subjects of intergovernmental fiscal relations at various levels [1, 2]. In this case, one has not been able to reach he desired hardness of budget constraints for the authorities of subjects of the Federation, to establish control over the efficient use of resources at the regional level, as well as to achieve the required growth formalization of intergovernmental fiscal relations [3, 4]. At present, fiscal regulation in Russia is overcentralized; therefore, many municipalities cannot function autonomously and sustainably, as local taxes and other local revenues make up less than 20% of their budgets. These problems cannot be solved without consolidating the municipal budget's local revenue base.

Russian municipalities differ noticeably in both the actual tax revenues and tax potential. In this respect, we can single out a group of municipal entities, e.g., the capital cities of oblasts and republics, i.e., centers of constituent entities whose financial statuses differ greatly from those of other Russian municipalities. The local self-government bodies of the constituent entities administrative centers, as a rule, have budgets comparable to those of the constituent entity itself (excluding the municipal budgets). We have selected the Novosibirsk oblast as a research target, as it

can be classified among the abovementioned group of Russian constituent entities. An analysis was conducted based on the data on the municipal districts (30) and towns (cities) subordinate to the authorities of the Novosibirsk oblast (5) in 2010-2016.

1 Specific features of revenue generation in the budgets of municipalities of the Novosibirsk oblast

The main revenues of local budgets are tax, non-tax revenues and grants from regional budget. Besides, it is only the tax revenues connected with economic potential of the given territory that can be regarded as a stable revenue base for the budgets of local self-government bodies [5]. To estimate the level of autonomy of local budgets, we have analyzed the distribution of the municipalities based on the share of collected (tax and nontax) revenues in the aggregate revenues of the local budgets. The calculations results are presented in Table 1.

Table 1: Distribution of the municipalities by the share of collected revenues, number of municipalities in the group

Share of collected (tax and nontax) revenues, %	2010	2011	2012	2013	2014	2015	2016
0-10	13	14	14	17	7	8	7
10-20	12	11	11	11	19	18	19
20-30	4	5	4	2	3	3	4
30-40	3	2	1	1	1	3	3
40-50	1	0	3	1	2	1	1
50-60	1	0	0	2	1	1	1
60-70	1	2	2	0	2	1	0
70-80	0	1	0	0	0	0	0
80 or more	0	0	0	1	0	0	0

As can be seen, the proportion of collected revenues for the majority of municipalities of the Novosibirsk oblast was in 2010-2016 within the limits of 20% . Thus, the situation has worsened compared to the end of 90th - beginning of the 2000s, when the share of this type of revenues for most municipalities of the Novosibirsk oblast was in the limits of 20-40% [6]. It is noteworthy that, over the period under study, in the Novosibirsk oblast, the proportion of collected revenues was more than 30% for a very small number of municipalities. Over 50% of collected revenues for the entire period were considered only in the city of Novosibirsk.

One of the weaknesses of the system of intergovernmental fiscal relations at the level of the subject of Federation is a high degree of centralization of budget revenues on sub-federal level, bias in favor of grants in the structure of municipal budget revenues [7, 8]. To test this assertion, consider the distribution of the share of grants in the aggregate budget revenues of municipalities (Table 2).

Table 2: Distribution of the municipalities by the share of grants in their budgets, number of municipalities in the group

Share of grants, %	2010	2011	2012	2013	2014	2015	2016
0-20	0	0	0	1	0	0	0
20-30	0	1	1	0	0	0	0
30-40	1	2	2	0	2	1	0
40-50	3	1	0	2	1	1	1
50-60	1	0	3	1	2	1	1
60-70	2	2	1	1	1	3	3
70-80	4	6	4	2	3	3	4
80-90	12	9	11	11	19	18	19
90 or more	12	14	14	17	7	8	7

It follows from Table 2 that, in the majority of municipalities of the Novosibirsk oblast, grants make up more than 70% of budget revenues, and consistently high throughout the period considered is the number of territories for which the share of grants exceeds 90% . In the structure of grants a high proportion of subventions and subsidies from the upper-level budget, which is caused by the transfer of the powers and financial resources from the regional to the local level.

2 Inhomogeneity characteristics of the fiscal capacity of municipalities

When comparing budgets of the same level, it is important to assess the expediency of concentrating resources from the standpoint of equalizing the municipalities' fiscal capacity and the levels of socio-economic development of the municipalities. This comparison can be conducted by using the per-capita inhomogeneity characteristics of the fiscal capacity before and after the municipal budgets were given grants from upper-level budgets [9]. We propose to use variation indicators as characteristics of inhomogeneity, i.e., the range of asymmetry, scatter, excess of scatter, standard deviation, and variation coefficient. With increasing homogeneity of the fiscal capacity in the sample, the variation indicators should go down. In our work we assessed the above indicators for the per capita collected and disposable budget revenues of municipalities (Tables 3, 4).

It follows from the data in Tables 3, 4 that the range of asymmetry between the municipalities in the Novosibirsk oblast after grant transfers from the oblast budget was decreasing in 2010–2015. The scatter of the municipalities based on the indicators of collected and disposable revenues, in general, increases over the period under study.

The excess of scatter is greater than 1 in the most of the explored cases. This is indicative that half of the municipalities with lower values of the indicators under study are close to one another in these indicators than the other half of the municipalities.

Table 3: Inhomogeneity of per-capita collected budget revenues

Indicator	2010	2011	2012	2013	2014	2015	2016
Range of asymmetry	8,34	12,22	5,08	6,70	5,81	4,19	3,28
Scatter	1915	2571	2244	2938	2609	2031	1541
Excess of scatter	1,128	1,325	1,254	1,327	1,264	1,237	1,148
Standard deviation	2598	4357	3182	4626	4211	2828	2210
Variation coefficient, %	57,71	84,76	57,17	67,85	57,02	42,87	33,21

Table 4: Inhomogeneity of per-capita disposable budget revenues

Indicator	2010	2011	2012	2013	2014	2015	2016
Range of asymmetry	6,13	4,00	4,15	3,16	3,11	3,72	3,49
Scatter	9461	7874	8742	10144	9366	10065	10387
Excess of scatter	1,124	1,000	1,038	1,023	0,998	1,044	1,052
Standard deviation	12219	11285	11954	13200	12252	12538	13085
Variation coefficient, %	39,17	33,39	32,13	27,37	27,71	30,27	29,90

The growth of the standard deviation of disposable revenue as compared to the standard deviation of the collected revenue is explained by the increase in the average level of the varied indicator. As seen from Table 3, the indicator of disposable revenues has a lower variation coefficient, i.e., the inhomogeneity of the municipalities' fiscal capacity after grant transfer from the oblast budget is decreased.

To determine which municipality groups experienced losses as a result of changes in the aggregate scatter indicators, we need to consider the changes in the distribution of territories based on the level of budget revenues as a result of money transfer from upper-level budgets. Tables 5-6 present the distribution of municipalities by the level of collected and disposable budget revenues per capita. Tables 7-8 shows the same, but centered values (the difference with the average for the region level).

Table 5: Distribution of municipalities by the level of collected revenues, number of municipalities in the group

Per-capita revenue, thousand rubles	2010	2011	2012	2013	2014	2015	2016
1-4	18	19	13	6	0	0	0
4-7	14	10	16	19	23	25	28
7-10	0	3	3	6	7	6	3
10-13	3	1	0	1	2	3	3
13 or more,	0	2	3	3	3	1	1

Table 6: Distribution of municipalities by the level of disposable revenues, number of municipalities in the group

Per-capita revenue, thousand rubles	2010	2011	2012	2013	2014	2015	2016
10-18	3	2	1	0	0	0	0
18-26	11	7	4	4	4	4	2
26-34	10	9	11	3	4	6	8
34-42	4	12	11	8	5	8	9
42-50	4	2	5	10	10	9	6
50-58	2	1	0	4	8	5	4
58-66	1	1	1	2	2	2	4
66 or more	0	1	2	4	2	1	2

The data in Tables 5-6 show that after the grant transfers to municipalities of the Novosibirsk oblast from the regional budget, there is sharp growth in the per-capita budget-revenue indicator by territory. If, before the transfers from oblast budget, the modal interval was from 1 to 7 thousand rubles of per-capita budget revenues, after the distribution of grants from the upper-level budget per-capita revenues increase dramatically. This indicates a significant increase in absolute and relative size of fiscal regulation resources in the municipal revenues in the Novosibirsk oblast. To exclude the effect of changes in the average level of budget revenues and assess the changes in their distribution with regard to the increased fiscal capacity standard, we have calculated centered values of the collected and disposable budget revenues.

Table 7: Distribution of municipalities by level of centered indicators of collected revenue, number of municipalities in the group

Per-capita revenue, thousand rubles	2010	2011	2012	2013	2014	2015	2016
Less than −2	8	13	10	16	10	3	2
−2-−1	8	7	10	5	12	17	11
−1-0	4	4	6	3	4	3	10
1-2	12	5	4	6	4	5	8
2-4	0	3	1	1	2	4	1
4-6	1	1	1	1	0	1	1
6-8	2	0	0	0	0	1	2
8 or more	1	2	3	3	3	1	0

If we take into account that all the municipalities of the Novosibirsk oblast are recipients of regional grants, which results in growth in the average level of fiscal capacity, then the outcomes of the oblast's fiscal policy appear to be less effective

Table 8: Distribution of municipalities by level of centered indicators of disposable revenue, number of municipalities in the group

Per-capita revenue, thousand rubles	2010	2011	2012	2013	2014	2015	2016
Less than −15	2	3	3	5	4	4	5
−15-−10	4	4	4	2	4	3	4
−10-−5	8	4	5	5	4	5	3
−5-0	6	6	8	7	5	6	8
0-5	6	10	7	6	6	5	4
5-10	1	3	2	3	6	5	4
10-15	4	2	3	2	2	2	2
15-20	1	1	0	0	1	3	2
20-25	1	0	1	3	2	1	1
25-30	1	1	0	2	1	0	1
30 or more	1	1	2	0	0	1	1

because there is an increase in both the number of territories with below average budget revenues and the number of municipalities with the highest revenues.

3 Assessment of the fiscal policy on the economic development of municipalities

In order to assess how well the current system of intergovernmental fiscal relations cope with functions of alignment of budgetary security differentiation of municipalities and encouraging municipalities to strengthen their own revenue base, you can use methods of regression analysis and ranking of municipalities in terms of collected and disposable budget revenues. With this interest are not the ranks, but changing them in the process of budgetary control. This change can be estimated by calculating the Spearman and Kendall correlation coefficients.

Spearman's rank correlation method allows to determine the closeness and direction of correlation between the two signs. Kendall's rank correlation coefficient determines the extent to which the ordering of all pairs of objects in two variables and is used to identify the relationship between quantitative and qualitative indicators, if they can be ranked. This ratio is preferable to calculate in the case of outliers. Values of rank correlation coefficients calculated for series of collected and disposable budget revenues are presented in Table 9.

As the calculations in 2011-2016 showed a moderate correlation between the ranks of collected and disposable revenues of municipalities' budget in the region. In 2010, the relationship was weak. Importantly, in 2010-2016 the relationship of analyzed signs was reversed, i.e. municipalities with large values of collected budget revenues had lower values of disposable revenues. This fact indicates that there is a significant

Table 9: Rank correlation coefficients

Coefficient	2010	2011	2012	2013	2014	2015	2016
Spearman's rank correlation coefficient	$-0,37$	$-0,47$	$-0,63$	$-0,44$	$-0,48$	$-0,48$	$-0,58$
Kendall's rank correlation coefficient	$-0,28$	$-0,37$	$-0,44$	$-0,33$	$-0,35$	$-0,34$	$-0,45$

change in the ranks of the municipalities of the Novosibirsk oblast after they received grants.

To answer the question of whether resources transferred from the oblast budget to local self-government bodies serve the purpose of intraregional equalization, it is of interest to determine the dependence between the grants from the oblast budget and per-capita tax or nontax local budget revenues. Therefore, we propose to estimate the following equation:

$$T_i = \alpha + \beta \cdot R_i + \xi_i,$$

where T_i are per-capita grants from the oblast budget to the ith municipality, R_i – indicates per capita tax and nontax revenues of the ith municipality, α is the intercept, β is the slope coefficient, and ξ_i are the regression residuals. The results of the calculations are presented in Table 10.

Table 10: Estimation results for the equation $T_i = \alpha + \beta \cdot R_i + \xi_i$

Indicator	2010	2011	2012	2013	2014	2015	2016
R^2	0,19	0,26	0,42	0,39	0,39	0,41	0,42
Estimate of α	36154	36126	47114	52399	52690	55877	65232
t statistics	8,88	12,03	12,82	13,92	13,37	11,63	10,78
lower bound of 95% confidence interval	27873	30015	39638	44743	41914	46104	52921
upper bound of 95% confidence interval	44435	42237	54590	60055	63467	65650	77543
Estimate of β	$-2,21$	$-1,52$	$-2,78$	$-2,09$	$-2,14$	$-3,19$	$-4,23$
t statistics	$-2,81$	$-3,38$	$-4,84$	$-4,56$	$-4,16$	$-4,76$	$-4,89$
lower bound of 95% confidence interval	$-3,81$	$-2,43$	$-3,95$	$-3,03$	$-3,42$	$-4,56$	$-5,99$
upper bound of 95% confidence interval	$-0,61$	$-0,60$	$-1,61$	$-1,16$	$-0,86$	$-1,83$	$-2,47$

The given data show that, in the Novosibirsk region during the period from 2010 to 2016 there was a statistically significant negative correlation between these pa-

rameters, that is the oblast fiscal policy is aimed at equalizing the per-capita budget revenues of the municipalities.

The literature has repeatedly emphasized that, in the given system of intergovernmental fiscal relations, local governments are not interested in implementing rational, transparent, or responsible fiscal policy. We can assess whether municipalities have positive or negative stimuli for responsible fiscal policy by the marginal effect of increases in taxes allocated to local budgets, i.e., by the growth of disposable revenue that results in the growth of tax revenues to the budget by 1 ruble, as follows:

$$(Y_{it} - Y_{it-1}) = \alpha + \beta\,(X_{it} - X_{it-1}) + \varepsilon_{it},$$

where Y_{it} – are disposable revenues of the ith municipality in year t, X_{it} – are tax revenues of the ith municipality in year t.

If there are no stimuli to increase tax and nontax revenues, then the regression coefficient β must be statistically insignificant. If stimuli (increase or decrease) are present, the regression coefficient shall be statistically significant (positive or negative). The city of Novosibirsk was excluded from the calculations. The presented results indicate that for 2010-2011, 2011-2012 and 2012-2013 years measured dependencies were found to be statistically significant; the estimate of the β coefficient in all the regressions is greater than zero. Thus, the stimuli worked towards conserving and developing municipalities' local tax potential. But other dependencies, namely 2013-2014, 2014-2015 and 2015-2016 years, were found to be statistically insignificant, that is transfers to municipalities were random.

Conclusions

An analysis of the revenue breakdown of the local budgets of the Novosibirsk oblast speaks of their low level of autonomy, since it is typical for them to not have any stable revenue base. At present, the proportion of collected revenues in the total sum of revenues in the local budgets of the municipalities of Novosibirsk oblast, is, on average, less than 20% . This means the dependence of the local budgets on the upper-level authorities.

This is supported by the recently increased centralization of the territorial budgets accompanied by an increase in the percentage of grants in the municipal budgets. In particular, in the majority of municipalities of Novosibirsk oblast, grants make up more than 70% f of all their revenues. As calculations have shown, this deprives the local self-government bodies of stimuli to fund their activities aimed at increasing their local tax base.

The system of intergovernmental fiscal relations is of economic, political, and social importance for the country's development. However, financial aid should play a secondary role in the development of a local tax base for budgets at each level. In order for the whole national budget system to function effectively, i.e., for budgets of different levels to be balanced and autonomous, it is necessary, first of all, to establish clear-cut and valid criteria for the distribution of tax revenues between budgets of all

levels.

Acknowledgements

The article was fulfilled according to the research plan of the IEIE of SB RAS within the framework of the priority direction XI.173 (Project XI.173.1.1, registration number 0325-2017-0004 in the ISGZ FANO, AAAA-A17-117022250125-4)

References

[1] Klistorin V.I. (2014). Federal relationships, regional policy, and the problem of deformation of Russia's economic space. *Regional Research of Russia.* Vol. **4**, Is. **4**, pp. 253-259.

[2] Suspitsyn S.A. (2017). Problems of coordinating macroeconomic and regional long-term solutions. *Regional Research of Russia.* Vol. **7**, No 1, pp. 62-70.

[3] Sumskaya T.V. (2017). The role of Decentralization in the Formation of the Budget System in a Federal State. *Economic Alternatives.* Is. **2**, pp. 293-306.

[4] Kolomak E.A. (2018). Spatial development of contemporary Russia. *Urban and regional resilience: strategies for success: RSA Russia conference 2018. 22nd-23 rd Oct. 2018. Regional Studies Association.* St Petersburg, 2018, pp. 9-10.

[5] Kolomak E.A. (2017). The post-Soviet evolution of the Russian urban system. *Area development and policy.* Vol. **2**, Is. **1**, pp. 24-39.

[6] Sumskaya T.V. (2009). Specific Features of Regional Budget Policy Using the Example of Novosibirsk Oblast. *Studies on Russian Economic Development.* Vol. **20**, No 4, pp. 374-382.

[7] Suspitsyn S.A. (2015). Measurements in the space of regional indicators: Methodology, techniques, and results. *Regional Research of Russia.* Vol. **5**, Is. **3**, pp. 223-235.

[8] Kolomak E.A. (2017). The post-Soviet evolution of the Russian urban system. *Area development and policy.* Vol. **2**, Is. **1**, pp. 24-39.

[9] Klistorin V.I. (2013). The Accuracy and Reliability of Forecasts. *Problems of Economic Transition.* Vol. **56**, No 01, pp. 35-40.

Data analysis in studying the geological section

Andrew G. Feldman, Natalia V. Molokova,
Dmitriy S. Rusin and Natalia V. Nikolaeva
Siberian Federal University, Krasnoyarsk, Russia
e-mail: anrfeldman@yandex.ru, nat_molokova@mail.ru,
rusin199812@mail.ru, nikolnat@mail.ru

Abstract

The article is devoted to developing a software module that allows interpreting field data, identifying locations of low and high resistance objects and evaluating them for the desired minerals.

Keywords: data analysis, geophysical data, interpretation.

Introduction

The rationale for this work is that interpreting received geophysical data has a high cost, long terms and inefficient outcome.

The goal of our work is to computerize interpreting data received with the long-wire method.

To achieve the stated goal, the following tasks were set:

- analyze the subject area;
- create a computer model;
- develop a software module;
- take into account ergonomic and psychological factors.

1 Source data preprocessing

Source data is field work materials taken with the long-wire method at the Barobinskiy exploration area in Kazakhstan (gold mineralization) by the alternating-current prospecting equipment developed by Krasnoyarsk geophysicists with contributions from the Chair of Radio-Electronic Systems of the Siberian Federal University [1]. Prospecting was carried out with a 3 km long wire rerouted each 1 km and a magnetic field survey from each wire position along orthogonal 1 km long profiles. The wire was powered by an alternating current of 312 Hz frequency. At the observation sites the horizontal component of the electromagnetic field H_y and the vertical H_z were measured.

The magnetic field intensity values, measured by the device, are total H^{Σ}, conditioned by the field of the wire H^{w} and anomalous objects H^{an}. It is possible to isolate anomalies at the observation profile only after subtracting the normal wire field from the measured values H^{Σ} : $H^{an} = H^{\Sigma} - H^{w}$. Due to the considerable complexity

of the theoretical definition of the normal wire field, this action is implemented by a practical technique in which the values H^{Σ} are averaged over all the profiles of $(\sum_{i=1}^{n} H_y)/n$. The effect of anomalous areas is minimized and the observation sites get anomalous values of the horizontal and vertical components at the observation sites - the formulas (1)-(2).

$$H_y^{an} = H_y - \frac{\sum_{i=1}^{n} H_y}{n}, \tag{1}$$

where n is a number of profiles; H_y is the measured value of the horizontal component.

$$H_z^{an} = H_z - \frac{\sum_{i=1}^{n} H_z}{n}, \tag{2}$$

where n is a number of profiles; H_z is the measured value of the vertical component.

As a matter of obtained anomalous values, further location analysis of the conducting objects is carried out. The analysis is based on the position [2], where the normals to the intensity vectors at the magnetic field observation sites indicate the place of its source, which is the region of eddy current concentration.

2 The computer model

The input parameters for the model are the measured values of the horizontal H_y and vertical H_z magnetic field components for pickets of various profiles and wire positions. The output parameters are the areas of occurrence of anomalous objects.

The full magnetic field vector is defined as the vector sum of the horizontal $\overrightarrow{H_y}$ and vertical $\overrightarrow{H_z}$ components with the coordinates $(H_y; 0)$ and $(0; H_z)$, respectively. As is known, if the vectors are given by their coordinates, then the sum of these vectors is a vector which coordinates are equal to the sum of the corresponding coordinates of the summand vectors. From here we find the formula for calculating the full vector coordinates - the formula (3).

$$\overrightarrow{H} = \overrightarrow{H_y} + \overrightarrow{H_z} = (H_y + 0; 0 + H_z) = (H_y; H_z)\,, \tag{3}$$

where H_y is the horizontal component value; H_z is the vertical component value.

Figure 1 shows the vectors for the pickets P_1 and P_2 located left/right of the wire position designated by the point O [2].

The next step is to plot normals passing through the observation points to the vectors.

The equation of line takes the form: $y = ax + b$, where a a is the slope of the line, b is some real number. To construct perpendiculars we use the following property - the product of the slope ratios of two perpendicular lines is -1. If $y = a_1x + b_1$ is the equation of the line which the vector lies along, and $y = a_2x + b_2$ is an equation of a perpendicular line, then $a_1 \cdot a_2 = -1$.

The equation of the line passing through two points $(x_1; y_1)$ and $(x_2; y_2)$ can be written as follows - the formula (4).

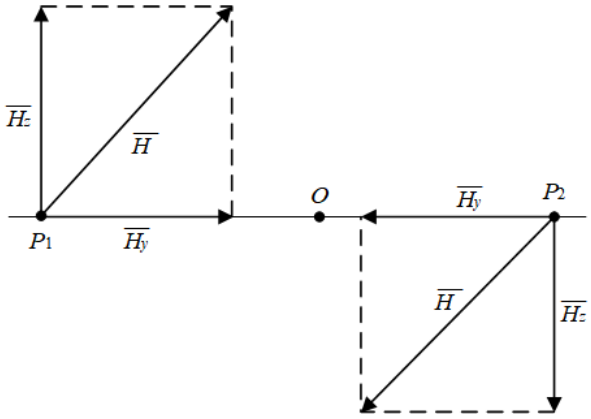


Figure 1: The full magnetic field vectors $\vec{H}$

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} \tag{4}$$

Having made the transformations, we obtain the equation for the slope ratio a_1 in the equation of line $y = a_1 x + b_1$, which the vector lies along - the formula (5).

$$a_1 = \frac{y_2 - y_1}{x_2 - x_1} \tag{5}$$

where $(x_1; y_1)$ – coordinates of the vector beginning; $(x_2; y_2)$ - coordinates of the end of the vector.

Then the slope ratio a_2 of the perpendicular line $y = a_2 x + b_2$ can be calculated by the formula (6). The ratio b_2 can be expressed by substituting the coordinates of the observation point $(x_0; y_0)$ through which the normal passes - the formula (7).

$$a_2 = -\frac{1}{a_1} \tag{6}$$

$$b_2 = y_0 - a_2 \cdot x_0 \tag{7}$$

To construct a perpendicular, we need to know its end coordinate. The length of the perpendicular is equal to the product of the coefficient of proportionality k the length of the vector by the length of this vector l. Figure 2 shows an example of finding the end coordinate for the perpendicular to the full picket vector located to the left of the O wire. Using the Pythagorean theorem, you can make the following equation: $\sqrt{(x_0 - x_1)^2 + (y_1 - y_0)^2} = kl$, where k is the coefficient of proportionality, l is the vector length, $(x_0; y_0)$ are the origin coordinates, $(x_1; y_1)$ are the end coordinates. Since $y_0 = 0$, and the coordinate y_1 can be expressed in terms of x_1 using the line perpendicular equation $y = a_2 x + b_2$, the equation can be converted to the following form $\sqrt{(x_0 - x_1)^2 + (a_2 x_1 + b_2)^2} - kl = 0$.

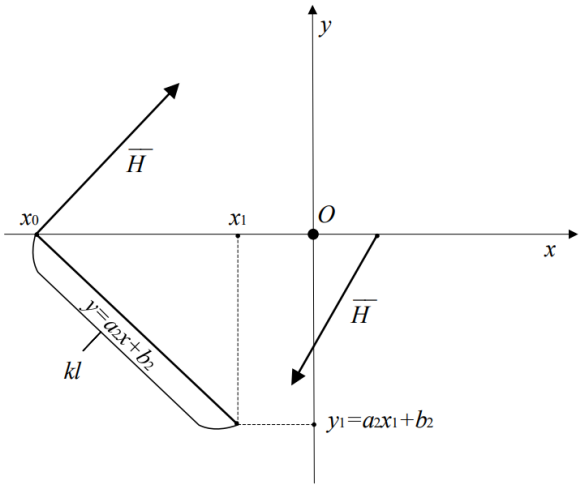


Figure 2: The example of finding the end coordinate for the perpendicular

The resulting equation is nonlinear, which can be solved using the bisection method.

The next step is to find the intersection points of the normals. To find the intersection point of two perpendiculars defined by the equations $y = a_1x + b_1$ and $y = a_2x + b_2$, we equate these equations $a_1x + b_1 = a_2x + b_2$ to the vectors and express x - the formula (8).

$$x = \frac{b_2 - b_1}{a_1 - a_2} \tag{8}$$

Substituting the obtained x value of the formula (8) into one of the two line equations that correspond to the perpendiculars, we obtain the coordinates $(x; y)$ of the intersection point.

You can find the concentration point areas using the k-means clustering algorithm with the Euclidean distance metrics. For two points A and B with the coordinates $(x_1; y_1)$ and $(x_2; y_2)$ the Euclidean distance is defined as follows - the formula (9).

$$d(A, B) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{9}$$

3 The problem algorithm

Predicting anomalous objects in the depth of the studied geological section on the field survey materials using the long-wire method is performed according to the algorithm:

1. At each observation point the anomalous values of the vertical H_z^{an} and horizontal H_y^{an} magnetic field components are defined by subtracting the wire field H_z^{av} and H_y^{av} from the observed values of the total field H_z^{ob} and H_y^{ob}.

2. The sum vector $\overrightarrow{H^{an}}$ is defined by vector summation based on the calculated anomalous values of $\overrightarrow{H_z^{an}}$ and $\overrightarrow{H_y^{an}}$.

3. The normal to the vector $\overrightarrow{H_y^{an}}$ is defined at the observation point.

4. The listed actions are performed at all the observation points of the profile.

5. Vectors and normals are constructed over the entire profile. The points of the normal intersections closest to the surface are secondary current concentration areas, i.e. anomalously conducting objects. The points of normal divergence correspond to the areas of high-resistance rock growth.

6. The anomalous object location is distinguished with the k-means clustering algorithm.

4 Software implementation

The software module is written in C #.

To store data we use the database, the scheme of which is shown in Figure 3. Microsoft SQL Server is chosen as a DBMS.

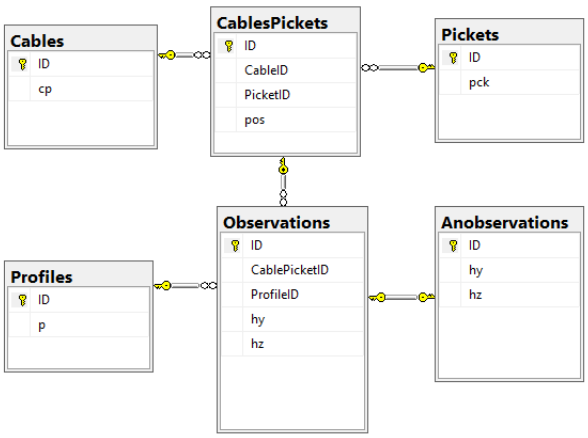


Figure 3: The database scheme

In the process, ergonomic and psychological factors were taken into account. A friendly user interface has been developed. For the convenience of geophysicists working with the software module, the created user guide comes with the description of operating with each module function.

Figure 4 shows the main window of the software module.

The interface allows a user to work with the database: add, modify, and delete records through the tab "Table". Figure 5 shows the form of interaction with the database for the "Observations" entity. In the DataGridView window, the observation values are displayed corresponding to the selected wire and picket position in the

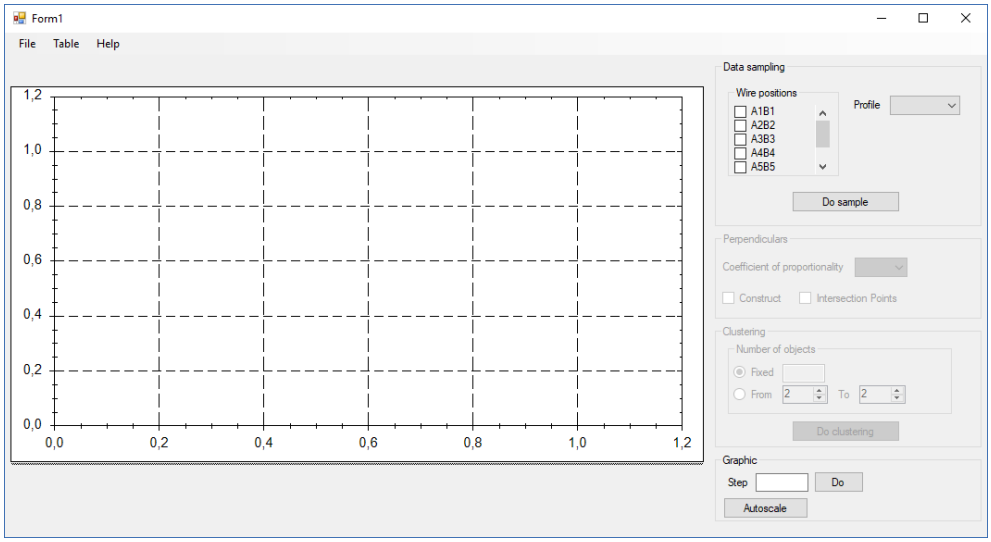


Figure 4: The main window of the module

ComboBox controls. The anomalous values of the horizontal and vertical components for observations are calculated after closing the form - the formulas (1)-(2) and added to the "Anobservations" table. The "Observations" and "Anobservations" tables are related one-to-one.

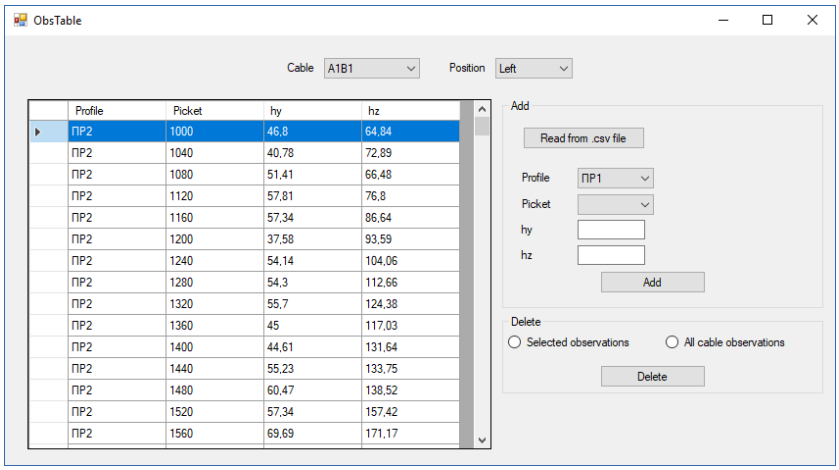


Figure 5: The "Observations" form

The right pane of the main window (Figure 4) contains the options available for selection.

Based on "Wire positions" and "Profile", the input data in the Sample () function is formed by sampling from the database. Figure 6 shows a chunk of the input data - the observations for Wire A1B1 of Profile 1.

After sampling the observations, we call the Vectors () function and construct full vectors.

	ID	CablePicketID	ProfileID	hy	hz
1	5025	66	1	163,91	115
1	5026	67	1	132,66	110
1	5027	68	1	142,11	115
1	5028	69	1	142,81	85,94
1	5029	70	1	110,16	95,94
1	5030	71	1	110,31	92,97
1	5031	72	1	80,39	81,88
1	5032	73	1	99,38	96,72

Figure 6: The chunk of input data

When a user selects the value of constant of proportionality, for each of perpendiculars the end coordinate is calculated. The coordinates are recalculated every time the user selects a new value.

Right after calculating, the "Construct" and "Intersection Points" elements become available for selection, and then the Perpendiculars () and Interpoints () functions are called, respectively. Perpendiculars to vectors are constructed with the Perpendiculars () function, and the Interpoints () function is used to find the intersection points of the perpendiculars.

Figure 7 shows an example of constructing the perpendiculars and their intersection points. The coefficient of proportionality is 3, wire A1B1, profile 2.

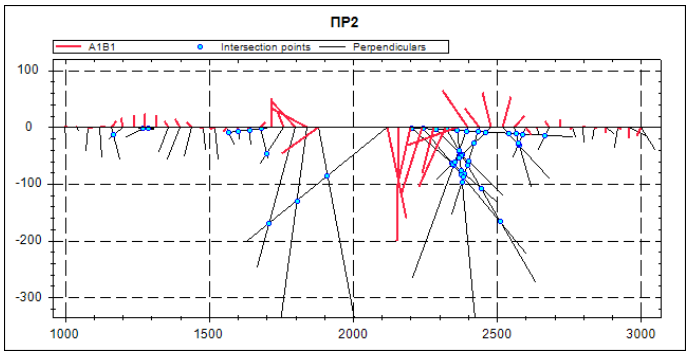


Figure 7: The perpendiculars to the full magnetic field vectors and their intersection points

To isolate concentration areas of the points that correspond to the predicted low-resistance and high-resistance objects when studying the geological section with the alternating-current long-wire method, the k-means clustering method was used. Figure 8 shows the program output.

The secondary field sources are concentration areas of eddy currents, arising in the development areas of highly conducting objects such as sulfide and magnetite ores or ground saline water. Areas without secondary electromagnetic field sources correspond to low-conducting rocks such as stone masses and gem stone deposits [3]. Each anomalous object is of interest for constructing a geological section. The geological nature of anomalous objects is determined with the assistance of auxiliary geological and mineralogical information.

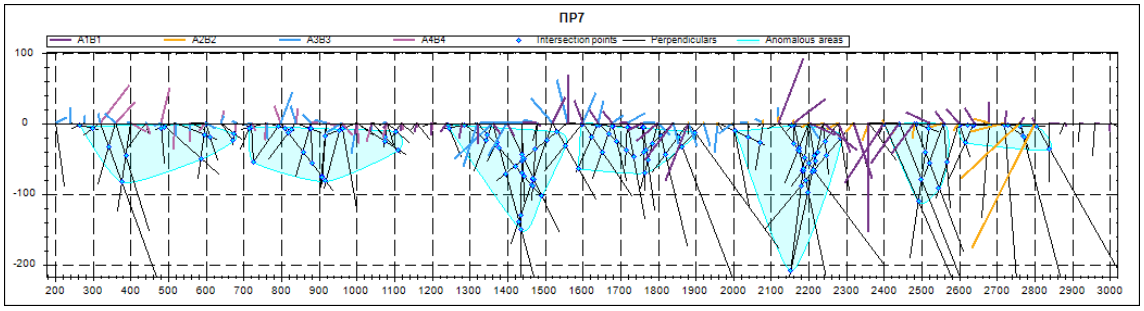


Figure 8: Predicting a conducting object in the geological section

Conclusions

The developed software module allows identifying locations of anomalous objects in the geological environment, to analyze the obtained locations of low or high electrical resistance, which is geophysical information as part of geological prospecting and exploration. In this case, the source information was processed and a geoelectric section was constructed for the Barobinskiy gold mineralization area. Computer-aided processing of large data volumes shows that there are low and high resistance areas in the section, which are of interest for prospecting and interpreting with the assistance of related G&G data. At the surface in the 800 - 1600 m spacing of the profile, fragments of quartz with gold were found, according to which it is assumed that we may find a series of gold-bearing quartz veins and stock work at the depth that correspond to high resistance areas. Low resistance areas may have sulfide mineralization. The eastern part has mostly near-surface highly conducting salt marsh. This software module helps significantly reduce the calculating time when processing materials. The processing tool increases transmission distance using the long-wire method and isolate anomalous zones. Goal-oriented well drilling of these zones allows reducing the financial and time expenditures of geological prospecting and exploration.

References

[1] V.I. Igolkin, G.Y. Shaydurov, O.A. Tronin, M.F. Khokhlov (2016). *Metody i apparatura elektrorazvedki na peremennom toke*. SibFU, Krasnoyarsk.

[2] A.I. Zaborovskiy (1963). *Geological prospecting (Elektrorazvedka)*. Gostoptekhizdat, Moscow.

[3] Feldman A.G., Sarvar S.S., Igolkin V.I. (2018). Estimation of the depth of the method of BDK by partial parameter first layer. *Problems of geology and subsoil development*. Vol. 1, pp. 453-454.

Computer-aided approach to synthesis the specialized frequency dictionaries

M. V. Karaseva[1],[2]
[1] *Reshetnev Siberian State Aerospace University 31, Krasnoyarsky Rabochy Av., Krasnoyarsk, Russian Federation*
[2] *Siberian Federal University, 79, Svobodny Av.,Krasnoyrsk, Russian Federation*
e-mail: karaseva-margarita@rambler.ru

Abstract

The practical use of the multilingual adaptive - training technology contributes to the intensive accumulation of specialized foreign terminology by students who study one or more foreign languages for their professional purposes. Electronic frequency dictionaries, built on a multilingual principle, are the basic components for supporting the multilingual adaptive – training technology. The article considers the computer-aided approach to the analysis of specialized frequency dictionaries.

Keywords: multilingual adaptive - training technology, frequency, dictionary, information basis, analysis.

Introduction

The purpose of the information multilingual adaptive - training technology is the intensive accumulation of specialized foreign terminology of students and specialists studying one or more foreign languages for their professional purposes. The main components of its support tools are electronic frequency dictionaries, built on a multilingual principle, and computer systems that implement the algorithm for training terminological vocabulary [1-5].

A frequency dictionary (information basis) is the information, whose portions are given to the student. The student will learn quicker and better the words and phrases that are found in the texts more often and these words are from narrow subject field. That is why it is necessary to take into account the frequency properties of texts. Table 1 shows an example of the information presented in the frequency dictionary [1]:

Eng	Ru
Decomposition algorithm, 49	Алгоритм разложения, блочный алгоритм, 15
Correlation analysis, 170	Корреляционный анализ, 110

Table 1: Portion of the frequency dictionary

Multilingual frequency dictionaries take into account the frequency properties of multilingual terminology [6]. An example of a portion of a multilingual dictionary on system analysis can be observed in Table 2

Eng	**De**	**Ru**	**Fr**
ability, 4	Fähigkeit, 4 f	способность, 8	Lacapacité, 5
above, 32	über, 100	выше, свыше, 19	Plushaut, 44
abstraction, 7	Abstraktion, 1 f	абстракция, 42	L'abstraction, 15

Table 2: Portion of the multilingual frequency dictionary

The dictionary in a broad sense is the most essential component of the speech perception model.A main operational unit is a word while speech perceiving. It follows that, in particular, every word of the perceived text should be identified with the corresponding unit of the student's (or a reader) internal vocabulary. It is natural to assume that from the very beginning the search is limited to some sub-fields of the dictionary. In a typical case the actual phonetic analysis of the sounding text gives only some partial information about the possible phonological appearance of the word, and this kind of information is obliged by a certain set of terminology; therefore, the problem arises (a) to select the corresponding set according to one or another parameter and (b) within the determined set (if it is adequately selected) to produce a "sifting out" of all words, except for the only one that corresponds best to the given word of the recognized text. One of the "drop out" strategies is the elimination of low-frequency words. It follows that the terminology for speech perception is a frequency dictionary. Partly for this reason, when the word identity is largely based on formal, graphical coincidence, semantics is taken into account insufficiently. In the result,frequency characteristics are biased and distorted. For example, if the words from the combination of "each other" are included by a compiler of the frequency dictionary in the general statistics of the use of the word "friend", then it is hardly to be justified. We are to recognize that in the word-combination we already have other words, and it is more precise that a separate vocabulary unit is a whole combination taking into account the semantics.That is why training is provided for both clearly formalized lexical units and well-established lexemes [4-6].

When organizing electronic frequency dictionaries that are the information base for supporting the multilingual adaptive-learning technology, the following objectives are pursued [2]:

- to reflect some of the important qualitative and quantitative aspects of general terminology on system analysis in English, German, French and Russian, resulting from statistical text analysis and description;

- to promote the organization of vocabulary learning and vocabulary accumulation in a rational way with the help of computer interactive training tools.

1 Texts analysis for the formation of electronic frequency dictionaries

The emerging directions of linguistics are connected with some applied tasks. At present, the direction of the applied linguistics has been practically formed. It is connected with the foreign languages training. Twenty years ago, a language description in teaching was considered as a problem of a methodology. Now it is clear that a methodology has other tasks, and a description of language phenomena should be carried out within the framework of linguistics and according to its laws.It is necessary to pay attention to the fact that the situation is similar to the way the direction of "computational linguistics" developed. Initially, problems connected with the automatic training were entirely attributed to the competence of programmers, and only then, the applied development of machine translation, automatic referencing, etc., was recognized as a linguistic value [3].

In the course of the language description some independent results were obtained. And, moreover, some developments that had appeared earlier, for example, functional grammar, linguistics of a coherent text, and possibly others, fell into the mainstream of this new direction.

Consider the peculiarities of pedagogical linguistics in connection with the problems facing this field.

The main peculiarity of the language description for the training purposes is to take into account the psychological properties of a person related to speech activity.These peculiarities are associated with the properties of memory (memorization, storage of information, its activation), understanding of speech and its generation and with peculiarities of communication, considered in a broader aspect (social, interpersonal relations, etc.). If we compare this psychological information with information about the structure of the language represented by the grammar, in many cases inconsistencies will be found. This fact is known to all teachers of a non-native language (especially if it is native to the teacher). First of all, they concern active speech activity since it requires the use of information on the compliance of language structures with the intentions of the speaker [6].

Language statistics can be defined as an auxiliary discipline of linguistics that explores the quantitative aspects of the language system use, including a professional-oriented system. Previously, some scientists used statistical methods successfully [2,3].Language statistics supplements the qualitative methods of language description through additional data characterizing the frequency of language phenomena. It is very useful in such practical field as information retrieval, lessons in foreign languages. Mathematically, this approach helps to model professional-oriented language communication as a probabilistic process, allowing one to determine the objective parameters of language differentiation expressed in various sublanguages, professional languages, professional dialects or styles.

Linguistics uses statistical methods, primarily in problems covering the language functionally, in texts, gathering separate passages into a coherent text. It is absolutely impossible to do in any other way because of the wide variety of language

communication in various professional fields.

When automating a general statistical analysis, the following steps can be distinguished: definition of statistical elements (word, phrase, sentence); determination of the absolute frequency of elements for a single sample and total sample; calculation of the relative frequency and probability for the main body of professional vocabulary terms; validation of the obtained frequency characteristics by calculating the standard deviations and the relative error; formalization of results in the form of lists, tables or graphs; interpretation and synthesis of results, up to the formulation of patterns.

Since it is almost impossible to cover the whole unity of subject-language communication even for only one language and one area, subject-language statistics should be based on the most representative sample tests, i.e. on the written or oral typical special texts. Each linguistic-statistical analysis begins with the selection and preparation of the corresponding text base. When specific tasks are set in the framework of applied linguistics, for example, when defining terminology for learning in a foreign language lesson or when compiling vocabulary for internal documentation, the text base can be very limited.

It is also necessary to pay attention to the type of the texts. Textbooks of higher and vocational schools of a general character are particularly suitable for the definition of a scientific and technical basic terminology. They guarantee a systematic, proportional and complete coverage of the material and the necessary language tools for its presentation.Moreover, they are less influenced by the individual language use by individual representatives of the profession. The further formation of the text base is built using new journals of a non-special nature.Reference books, reports, progressive messages, and instructions for application other types of texts, by contrast, are a favor able starting point for observing subject-language peculiarities at the level of both a sentence and a text.

The first result of statistical text processing is absolute frequency. It shows how often the corresponding phenomenon occurs in the text under study. However, it has little value for further research in the practical use of the results or in general for generalized statements, since it directly depends on the size of the selected text.It serves mainly as an initial value, for example, for calculating the relative frequency.

The relative frequency is a percentage that expresses the proportion of a language unit in the whole text.It is obtained by dividing the absolute frequency by the length of the sample, for example, for a word with a particular 186 in one sample from N = 50,000, the relative frequency will be calculated as 186 / 50,000 = 0.00372.

In other words, the relative frequency of a phenomenon is a ratio of the number of its actual occurrence to the number of its theoretically possible occurrence.It is possible to equate the relative frequency to the probability of a linguistic phenomenon if a sample is representative for a subject language.Then it gives grounds for statements about the statistical structure of the relevant sublanguage or about the importance of individual elements for the text organization.

A particularly important step in the linguistic and stylistic analysis is the reliability control of the determined data. There exist various ways of control.The standard

deviations (errors), the relative error, and the confidence interval are primarily taken into account in stylistic statistics and subject-language statistics.

Standard error (mean square error) is a measure of the variability of the average frequency of a linguistic phenomenon in partial sampling. Its calculation is done according to the formula:

$$S = \sqrt{\frac{SAQ}{n-1}}, \tag{1}$$

where S is a standard error; SAQ is a sum of the square of the error; n is a number of control samples.

The relative error is calculated for certain lexical units in frequency dictionaries to determine the accuracy of these dictionaries. It is done according to the formula:

$$|f - p| = Zp\sqrt{\frac{p(1-p)}{n}}, \tag{2}$$

where f is a relative frequency; p is a probability; Zp is a coefficient for a given level of confidence p; n is a volume of the sample control.

In papers concerning with the linguistic, the simplified versions of this formula are used. They assume that the difference (1-p) is approximately equal to one for a small p. A general variant of determining the relative error is as follows:

$$\delta = \frac{Zp}{\sqrt{nf}} \text{ or } \delta = \frac{Zp}{\sqrt{F}} \tag{3}$$

where δ is a relative error; Zp is a coefficient for the given level of confidence p; n is a sample size; f is relative frequency; F is absolute frequency.

The calculation of the confidence interval is a sophisticated version of the calculation of the relative error.The lower and upper limits (p_1 and p_2) of the oscillations and the average frequency are determined. There exist different methods of calculation, for example:

$$p_1 = \frac{fN + \frac{1}{2}Zp^2 - Zp\sqrt{f(1-f)N + \frac{1}{4}Zp^2}}{n + Zp^2} \tag{4}$$

$$p_2 = \frac{fN + \frac{1}{2}Zp^2 + Zp\sqrt{f(1-f)N + \frac{1}{4}Zp^2}}{n + Zp^2} \tag{5}$$

χ^2 - test, determines whether any observed differences occur in different samples, or whether the samples belong to the same basic population (functional style, sublingual, objective language, type of text, etc.). In most cases, this involves the verification (authenticating or falsifying) of the main assumption (null hypothesis); for example, the expectation that word types play a similar role in the text.The reference value χ is a sum of the observed and expected frequencies for a certain number of variables referred to the expected frequencies

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_{bi} - f_{ei})^2}{f_{ei}}, \tag{6}$$

where k is a number of variables; i is a variable; f_{bi} is the observed frequency of the variable; f_{ei} is the expected frequency of the variable.

When comparing samples, the expected frequency of f_{ei} is usually equated to the average frequency $\overline{x}$.

$$\chi^2 = \frac{(x_i - \overline{x})^2}{\overline{x}} \tag{7}$$

The subject-language statistics use various lists, tables, and graphs to present the research results.With the help of a circular image and strip charts first of all, parts in percentage values are given. Histograms and a chain of polygons are suitable for the graphic representation of quantitative peculiarities, such as word length or sentences length. Curves with a relatively typical flow exceed this simple combination of frequencies in qualitative and quantitative terms. They help to recognize the functional relationships between signs and their frequency.The frequency of language phenomena can itself become a sign characterized by other data.

There exist, for example, the following dependencies:

- between the frequencies of lexical units and their classes in the frequency dictionary;
- between the frequency and its probability of occurrence in the text;
- between the frequency and a relative error;
- between classes of one frequency dictionary and the cumulative number of lexical units;
- between classes and class wrappers;
- between the frequency and potency of communication;
- between frequency and degree of specialization of the subject lexical vocabulary;
- between the length of the text and the amount of vocabulary, etc.

Linguo statistics helps to determine which language phenomena occur in speech or texts more often.Statistical methods study the vocabulary of the language intensively. Frequency dictionaries give the information about the general vocabulary.A frequency dictionary registers words, word forms or word-combinations that have been encountered in the text (sample) studied for its compilation. Together with these units (i.e., words, word forms, word combinations), their frequencies are indicated in the dictionary, i.e. numbers show how many times each dictionary item has been encountered in the given text [6].

Conclusion

As it has been already noted, the frequency dictionary indicates a number of cases of the word use in texts that were analyzed to compile a dictionary. Frequency dictionaries differ depending on the principle of position of the database position. Words or word-combinations can be placed alphabetically, as in a typical dictionary, with its frequency to a word. Also, words and word-combinations can be arranged in descending order of frequencies, starting from the most commonly used word. The first version of the dictionary is addressed to a student; the second one is to the learner. A student can also work with the second version of the dictionary to learn a foreign language independently, for example, when memorizing words and word-combinations in portions, depending on their frequency or checking proficiency in terminology units, starting with the most frequent ones.

References

[1] Kovalev I V, Karaseva M V 2013 *English-German-Russian frequency dictionary on systems analysis in electronic engineering and aerospace*(SibSAU Krasnoyarsk)p 216

[2] Kromer V V 1997 *Nuclear Fan Model of Vertical Word Distribution in Russian*(Dep. in INION RASNSPU Novosibirsk) N 52458 pp 132–146

[3] Mueller Ch 1968 *Initiation a la statistiquelinguistique* (Larousse) p 249

[4] Kovalev I V, Kovalev D I, Karaseva M V, Pershakova K K, Tueva E V 2017 *Multilingual environment of information and educational interaction* (Scientific and technical information Series 2: Information processes and systems) N 7 pp 24–31.

[5] Engel E A, Kovalev I V and Engel N E 2016 *Control of technical object on the basis of the multi-agent system with neuroevolution and student-teacher of-line learning* (IOP Conf. Series: Materials Science and Engineering) 155 012001

[6] M. Karseva Informational and Educational Interaction for Multilingual Environment. Applied Methods of Statistical Analysis. Nonparametric methods in cybernetics and system analysis – AMSA'2017, Krasnoyarsk, Russia, 18-22 September 2017: Proceedings of the International Workshop. - Novosibirsk: NSTU publisher, p.350-355.

The income prediction module of the retail store's network

Kristina Pakhomova[1], Pavel Peresunko[2], Sergey Videnin[3] and Eugenia Soroka[4]
Institute of Space and Information Technology, Sibirian Federal University. Russia
e-mail: [1]kpakhomova@sfu-kras.ru,[2]peres94@yandex.ru,
[3]svidenin@sfu-kras.ru,[4]jollot@yandex.ru

Abstract

The main idea of this paper focused on the development of a program module, which predicts the pharmacy retail income by the machine learning theory. Beyond that, we want to introduce the best prediction model, which has learned by specific retail dataset. Notice, the architecture of program involves dynamic upload dataset, by Yandex" Internet service. The dataset represents the set of features and set of retail points, however in this task, the features describe the pharmacy industry. So on the first step will analyze the dataset and found out the correlation of the features. Next, will select the relevant features, which affect on income rate of the retail point. The last one will introduce to the prediction income Average model. In the last, will compare the three models, there is Average model, Gradient Boosting Regression and Random Forest Regression.

Keywords: Feature selection, Income prediction, Machine Learning, Artificial Intelligence

Introduction

Nowadays, Artificial Intelligence tools have become more actual in a business area, particularly machine learning approaches provide insightful economic analysis and an increase of company income. Indeed, business needs on the statistics, economics, marketing, and mathematical approach, for the analysis of the important retail features and profit markup. Sometimes the businessmen have not sufficient information about the features quality, which affects the profitability of the retail stores. Traditionally, the efficiency of features was calculated or analyzed by experience way. However, in the context of selecting the new retail store with an estimated future income rate, the businessman may make mistake. The experts need to aggregate a stores information (humans traffic, income rate, count of bus or metro station, business competition and etc) for drilling down a database of features. So, identifying the important and relevant features has a number of an advantage today.

Actually, machine learning provides a variety of choice of universal algorithms focused on feature selection, recommendation, prediction, and other tasks. Anyway, those algorithms do not provide the reliable result in a case a small sampling of fewer than 100 units and features count more than 120 units. Previous, the task of prediction income of retail was investigated by [5, 3]. By the way, the most advanced

method is neural network [1], where authors introduce the simple linear regression then multiple regression and after describe the algorithm of neural network, which applied on a retail dataset from "Google Places API". In this paper we want to introduce our decision for prediction retail store income and compare it with the general machine learning methods.

This paper introduces the module for the retail store's income prediction based on the mathematical model with the specific dataset. In section 2, we will describe the prediction algorithms and introduce the Average method. In section 3, we will describe the dataset, particularly the set of features and their multiplicity. By the way, the huge number of features according to their relevant selection is scaling each feature to a given range. In the last section, we will apply the method to the dataset and compare the Average method with other general prediction algorithms.

1 The module development

Every store in the retail network has a set of features which influence on the successfulness of the store. By the way, for each retail store network has specific particular qualities describe an income of store. In this paper we want to introduce the module, which analyses the features of the stores' network, select only relevant features and predict the income for a future stores for target retail network.

For this reason, the Figure 1 shows the general mechanism to this paper, where we will introduce the prediction module, which runs a red contour. For the experiment the dataset was obtained by "Yandex" services and custumer's statistical informational.

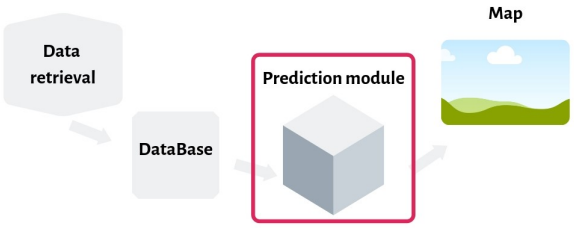


Figure 1: The diagram of the problem conception

The Prediction module includes the main three step, where on the first we normalize the dataset, then selected the relevant features by Boruta alghoritm and in the end to choose the best prediction alghorithm for set a task.

The raw data should be uniform in the case of the heterogeneous dataset. Indeed, exist a several type of features, where a dataset may have a binary type, counter and value. For this reason, we apply the min-max scaling procedure in the interval [0,1].

On the second step, we want to reduce the number of features and compute only relevant, which influence the income rate. In this experiment, we will apply the

Boruta algorithm, where the features are copied, and then each new feature is filled randomly by shuffling its values. Random Forest is started on the resulting sample. In order to obtain statistically significant results, this procedure is repeated several times, variables are generated independently at each iteration.

On the third step, we apply the prediction algorithm to dataset except for irrelevant features. In the experiment we tested 3 methods, there are Gradient Boosting Regression, Random Fores Regressor, Ridge regression, Average method. During experiment we found out the tree-like structure algorithms is a set of decision's tree, which provide the best result of the prediction. Given a traning set $X = x_1, ..., x_n$ with responses income for each shop $Y = y_1, ..., y_n$, where N the count of shops. The Random Forest describes as:

$$\hat{Fr} = \frac{1}{B}\sum_{b=1}^{B} fr_b(x^{'})$$

where B is the number of trees and predictions for unseen samples $x^{'}$ can be made by averaging the predictions from all the individual regression trees on $x^{'}$.

The Gradient Boosting builds the model in a stage-wise fashion like other boosting methods do, and it generalizes them by allowing optimization of an arbitrary differentiable loss function [6].

$$\hat{Fg} = argmin\mathbb{E}_{x,y}[L(y, F(x))]$$

where $L(y, F(x))$ is loss function, which describes $L(y, F(x)) = (F(x) - y)$ so $\frac{-\partial L}{\partial F(x)} = 2(y - F(x))$. The parametr α has choosen 0.9, maximal depth = 3,the number of boosting stages to perform = 100.

Howewer, our idea was the join in two Bagging and Boosting algorithms together that in this paper introduced as Average method. Indeed, we average the prediction result after test sample of two methods: Gradient Boosting Regression and Random Forest Regression.

$$\hat{Fav}_i = \frac{\hat{Fg}_i + \hat{Fr}_i}{2}, i = 0...22$$

The Regression methods calculate the value of prediction more inaccuracy, it proves the MAE and MAPE estimation in the fourth sections. Those algorithms are the basic tools of the statistics and them implementation was introduced in the paper [4]. For this reason, we did not introduce more information about them.

In this section we introduced the general prediction algorithms in machine learning. Also, we described the theoretical part of the retail location prediction module. On the next section we will focus on the retail store's features.

2 The retail features introduction

At the beginning, we want to consider the retail stores features and their influence on a store income rate. At the end of this section, we will introduce the features correlation heatmap.

Nowadays, the feature selection has been proccesed by general machine learning algorithms. They combine the computational complexity, the quick operation and universal, consequently, the choice of optimal algorithm depends on a formulation of the problem and the initial dataset. There are a few papers, which describe the retail feature analysis [2, 3]. One approach focuses on prediction of the retail location by using neural network [1]. We presume that an accuracy of prediction depends on a set of features, which influence income rate.

Indeed, each store has a number of features, which are usually common to all stores belonging to the same retail network. Thus, each feature influences the income rate with a greater or lesser degree, accordingly, the owner may manage to income by establishing the necessary measures of the features. Actually, there are no specific features that clearly affect a profit, because the number of important features depends on the category of the store network and the city urbanization. For example, in the case of pharmacy retail, an important feature is the presence of hospitals and medical institutions close to the store. We note that there are common features for all store chains, such as the area of the premises, the number of cash desks and sellers, the availability of parking, traffic, bus stations and others.

The paper will investigate the pharmacy industry, where the retail chain has the set of stores in the count of 22 items and each store has features in the count of 134 items.

Actually, the pharmacy industry has several specific types of the features. Some of them have a radially distance dependence on the store location. An example might be the count of bus station from the store at a distance of 100 meters, 200 meters, 300 meters, 400 meters, 500 meters and 800 meters radially. Another features is square of the store in the square meters. Table 1 describes a few features of the pharmacy industry.

Moreover, the heatmap of correlation matrix on a Figure 2 shows the dependence between several features and income. The deep red color identifies a strong positive correlation, but the deep-blue is a negative correlation. For instance, the feature "rubric 365-800" means a number of the hospital located within a radius of 800 meters has a strong correlation with the feature "pharmacies-800" - the number of competition on radius 800 meters.

In the next section, we will refer to the implementation of the prediction model on the pharmacy retail data set.

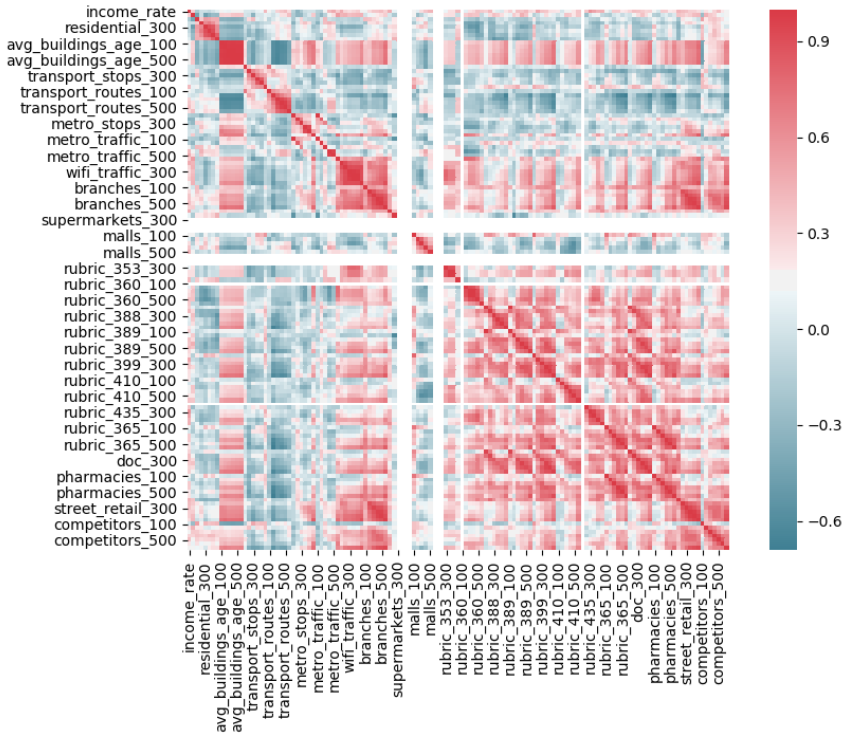


Figure 2: The correlation matrix of the retail features

3 The module prediction on an experimental retail data

The prediction module consists of the following steps,

1. The retail data preprocessing includes the scaling procedure. In brief, we use the min-max scaling, which consists in rescaling the range of features to scale in the unit interval. In addition, the profitable to ranking the stores on income rate, where the first shop the low income and last the high income, to observe the income fluctuation.

2. Indeed, several features did not corelate to income and the count of featuers are massive that uninformative. For this reason, we need to select informative features, which correlated to income. Therefore, in this experiment, we apply two-step of selection, where on the first, the decision maker selects irrelevant features or noisy dataset. In the second, we apply feature selection algorithm Boruta, it works as a wrapper algorithm around Random Forest.

3. Next, we apply the prediction Average method to dataset after preprocessing procedure. Indeed, in the comparison diagram, we may observe the advantage of this method.

Table 1: The retail data before and after scaling procedure

storeID	area/scaling	residental 100/scaling	metro stops 200/scaling	malls 400/scaling
1	79.6/0.4688	331/0.4203	1/0.5	0/0
2	93.9/0.5750	513/0.7046	2/1	1/1
3	70.1/0.3982	432/0.5781	1/0.5	0/0
4	77/0.4495	293/0.3609	0/0	1/1
5	80.4/0.4747	702/1	0/0	1/1
22	101.7/0.633	288/0.3531	0/0	1/1

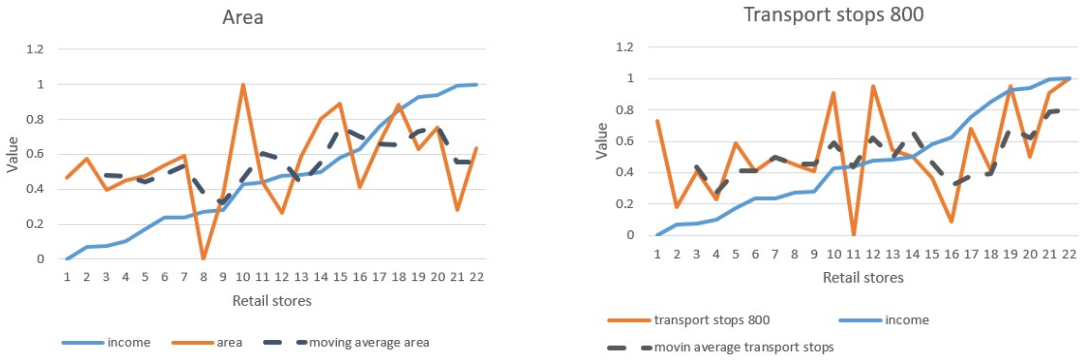
4. In the end, we build the project as a general module for income prediction. After, this module will integrate to the main business logic for the retail network system.

Certainly, before the avareged method implementation, we need to organize the retail store data preprocessing procedure. The experimental sample has 22 stores and 134 features. The experiment was implemented by using a Python 3.4 version. The dataset was given by the medicine company and was supplemented by Yandex services.

At the first, we need to structure information by ranking the stores according to the income rate. At the second, we implement the preprocessing procedure as the scaling of data set. It was computing by using min-max approach, which provides distribution values on the unit interval. The intermediate calculation was presented in Table 1.

After, we apply the Boruta algorithm to normalise dataset. Boruta is heuristic algorithm for selection of significant features based on the use of Random Forest. It tries to capture all the significant features in the dataset with respect to an outcome variable in our case is income. Therefore we defined 14 features, which depend on income change. There are "area", "transport stops 800", "metro traffic 400-500" and "metro traffic 800", "wifi traffic 100", "street retail 200 - 300" and etc. Indeed, Figure 3 describes two features "area" and "transport stops 800". On the first figure we observe the dependence on store's area and them income in the interval between 0.3 and 0.8. The moving average line shows the trend line for feature "area". The tails of moving the average line in interval [0;0.3] and [0.8;1] describes the low dependency on the store's income. The second figure shows the simillar case, wherein the interval [0.4;0.7] we observe the strong dependence on numbers of transport stations to radius 800 meters to the store's income.

According to this number of features, we apply the prediction algorithms to the preprocessing dataset. From a large number of prediction algorithms, we have choosen the Gradient Boosting Regression, Random Forest Regression and Avarage method. The tree structure algorithms provide more simple interpretation, flexibility, and high computational speed the specific dataset regardless of the count data and the variety. The results of the data prediction was provided by Table 2. Statistical estimates are used to estimate the prediction, where the simplest is the deviation of the fact



a) Feature "area"

b) Feature "transport stops 800"

Figure 3: The dependence between income and features "area", "transport stops 800"

Table 2: Comparison the estimations of prediction algorithms

	MAE	MAPE
Gradient Boosting Regressor	0.29663477	142.6556371
Random Forest Regressor	0.261513839	123.5678737
Ridge	0.567570237	220.6778885
Average method	0.273655219	131.8967422

from the prediction in quantitative terms. The value of mean absolute error supports Random Forest Regressor.

By the way, the dataset was devided on traning set and testing, but the number of samples was so small and was difficult for algorithms predict the exact result. So, we trained algorithms except for one store and we tested only this store. Other words, all stores are 22 iterations, where we except one store trained by 21 stores, where one was testing. During all iterations, we excepted each store from learning. Figure 4 shows the three prediction methods, at first thought all methods are simillar but the curve of Random Forest Regressor more flat than other and it MAE value is low. However, it is a fact, that the number of samples was poor and the algorithms do not provide exact result of the prediction, because they necessary the much more quantity of sample. This problem we may observe on the tails of curve, where the low-income stores has only 4 units and a similar situation with the high-income stores.

In summery, we computed all steps for explaining the Prediction module which was included on the union project for the pharmacy retail store's network. By using this program common person may analyze the location on a city map according to the features and follow to this prediction's advice.

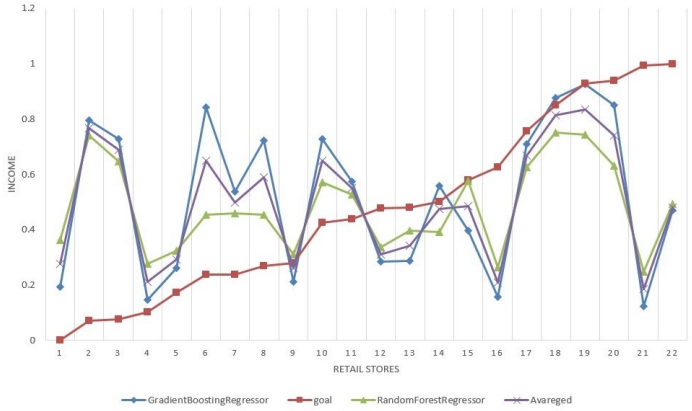


Figure 4: The comparative diagram of prediction store's network income

Conclusion

In this paper, we focused on a pressing task as a prediction income of the retail chain. The prediction module based on Average model, which is union the Gradient Boosting Regression and Random Forest methods, the feature selection was computed by Boruta method. Currently, this module would be extended not only the pharmacy industry but also on the other business area. Despite it, this module was included in the global service "Mestomer".

References

[1] Satman M.H., Altunbey M. (2014). Selecting Location of Retail Stores Using Artificial Neural Networks and Google Places API. *International Journal of Statistics and Probability*. Vol. **3**, pp. 67-77.

[2] Ferreira K., Lee B.H., Simchi-Levi D. (2016). Analytics for an Online Retailer: Demand Forecasting and Price Optimization. *Manufacturing & Service Operations Management*. Vol. **18**, pp. 69-88.

[3] Glaeser C.K., Fisher M., Su X. (2016). Optimal Retail Location: Empirical Methodology and Application to Practice. *SSRN Electronic Journal*. pp. 1-28.

[4] Duda R.O., Hart P.E., Stork D.G. (2000). *Pattern Classification (2Nd Edition)*. Wiley-Interscience, New York, NY, USA.

[5] Cheng E., Li H., Yu L. (2007). A GIS approach to shopping mall location selection. *Building and Environment*. Vol. **42**, pp. 884-892.

[6] Friedman J. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*. Vol. **29**.

Research of educational business processes in the decision making support system of University

Vladimir M. Stasyshin
Novosibirsk State Technical University, Novosibirsk, Russia
e-mail: Stasyshin@ciu.nstu.ru

Abstract

Some directions and results of research in the field of educational data analysis using the NSTU Information System are presented. Such results are useful for administrative decision-making to university business processes. Received results were discussed at various levels of administrative management of the University and served as the basis for the improvement of educational business processes at NSTU in the Decision Making Support System.

Keywords: information system, educational data mining, data analysis, decision-making support, classification algorithms.

Introduction

Universities are operating today in a very dynamic and competition environment. The rapid development of information and communication technologies have led to very strong competition between universities. Leading educational institutions have realized that they need to be proactive, introducing innovative management approaches and using progressive methods and techniques, in order to remain competitive [1]. Solving the problem of raising a level of higher education, the state places higher educational institutions in conditions of competitive struggle, in which only universities that have managed to adjust according to the demands of the time and mobilize all their resources for this purpose win. The goal pursued by the state is to get rid of weak universities in the market of educational services. This requires the administration of the university to make timely management decisions to improve the effectiveness and educational process and other areas.

The decision-making process in the field of education is very complex, as it involves a wide range of stakeholders who are related to higher education in one way or another (teachers, employees, students, their parents, ministry employees, etc.). Inevitably, this process contains many factors of a subjective nature. Among other things, a competent analysis of information accumulated in the university's information system is one of the elements to make it possible to increase the level of objectivity and quality of decisions. University management have to take quickly important decisions, and therefore information are needed timely and high quality.

Modern universities are collecting large volumes of data referring to their students, the organization and management of the educational process, and other business processes. Huge amounts of data have accumulated on all aspects of the vocational guidance, educational, financial, economic, scientific and other spheres of the university's

activity in the Information System (IS) of the Novosibirsk State Technical University (NSTU) since 15 years of its development. This data may be use for making management decisions to improve the business processes of the university [4]. The available data usually used for producing simple and traditional reports. Moreover, much of the data remains unused due to the inability of the university administration to handle it because of the large volumes and the increasing complexity.

Analysis of educational data (EDM - Educational data mining) for improving business processes in the university is one of the modern trends in the education system. It's necessary to introduce advanced information technologies to effectively transform available data into knowledge to support decision making.

This paper describes the results of researches in the next areas:

- the analysis of the correlation of the USE (Unified State Examination) score and the progress of students,
- the studying on factors affecting the students' expulsion,
- the analysis of the demand for educational resources in the Electronic library system of NSTU,
- the research of approaches to improving the quality of higher education, carried out using Methods of Expert Estimations.

1 Analysis of the correlation of the USE score and the of students' progress

It is considered that the average USE score at admission is one of the indicators of the «effectiveness» of the university. This is one of the indicators that the university reports to the ministry. The generally accepted qualitative model of dependence of the student's progress at the university of his USE score (Model 1) has the form shown in Fig. 1.

The research task was to confirm or disprove the statement that this model is the only true one [5]. Research of performance by semester and USE average scores of entrants performed using developed software for the storage and delivery of reports and tools for regression models building from Caret package of software system R. Analysis of the reports and regression models showed the dependence of academic performance of USE average score doesn't always have the type of dependence presented by Model 1. Students not having the highest score of the exam often show better academic performance. This type of dependence (Model 2) presented in Fig. 2. Despite the fact that the Model 1 is dominant, and its nature is observed in most of the cases, the Model 2 of academic performance of USE average score have place in a number of educational programs in some departments, for some semesters in some faculties [5].

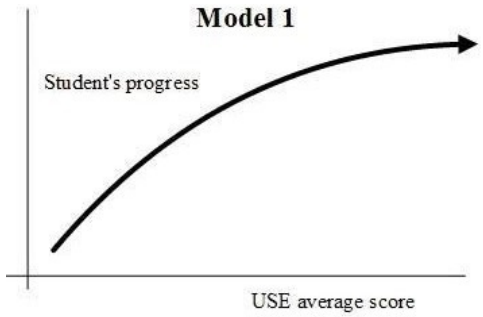


Figure 1: The generally accepted qualitative model of the student's progress in the university dependence on his USE average score

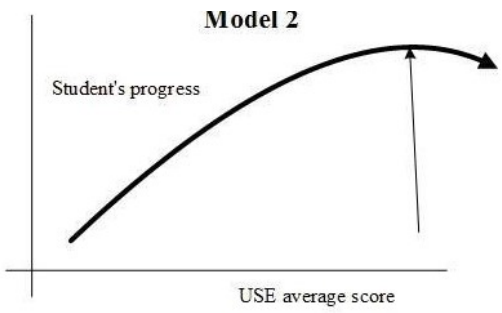


Figure 2: The alternative qualitative model of the student's progress in the university dependence on his USE average score

Below is a regression model of student academic performance of direction 02.03.03 «Mathematical support and administration of information systems» (data for 2011-2017), its high-quality appearance corresponds to a Model 2 (Y – performance in the 100-point system, X – USE score):

$$Y = -4.3334 + 0.2263X - 0.0015X^2 \quad (1)$$

In general, it can be concluded that the USE average score can not be the basis for the definition of the university's «efficiency» in terms of «education quality» in the sense of student performance. Some teachers who were familiar with the presented data suggested that the nature of Model 2 is a sign of certain problems in the study process: congestion of students, imbalance of training courses, etc.

2 Research is the investigation of factors affecting the students' expulsion

The task of retaining students is the classical task of EDM, and many publications are devoted to its solution, for example [2,3,6]. Like any university, Novosibirsk State Technical University is extremely interested in retaining of entered students, so in advance it would be useful to know students from «at-risk group» with high probability to be expelled in order to take up to date measures and to provide assistance to students (special training conditions, additional consultations, etc.). The research aim is to find out whether there are templates (sets of input variables-factors) that can be useful for predicting student expulsions (a variable of the output class). Analyzed factors in terms of their impact on students' expulsion are:

- an average USE score on admission;
- a category of secondary school (lyceum/gymnasium /secondary school of the regional center/district secondary school/ rural secondary school/college/secondary technical school);

- place of residence (at home/in a residence hall/ at an apartment);

- social welfare (satisfactory/low: certificate from the department of social protection, orphan, disabled);

- enrollment (on a general basis/target set);

- average score in the first session.

The research used data from the University's Information System: data for 2011-2014 are used to train the classifier, data for 2011-2018 are used for testing. The analysis was carried out for all faculties and for all educational programs. Several different algorithms were used from different packages of the software system R to build the classification model: decision tree algorithm Cart (function Rpart from Rpart package), random forest algorithm (function RandomForest from RandomForest package), Support Vector Machine algorithm (function Svm from e1071 package). Each classifier was used to test two variants: Variant1 with cross validation by dividing the original sample into 10 equal parts (9 parts were used for training and one – for testing) and Variant2 – the percentage of separation (2/3 dataset used for training and 1/3 – for testing). For the majority of faculties and educational programs, a high accuracy of the classifier was achieved at the level of 75-82%. Some results of research are given in Tab. 1.

The classification trees obtained by the CART classifier contain from one to five levels of hierarchy. The attribute "Average score in the first session" appears at the first level of the tree for all faculties, for such faculties as Faculty of Automation and Computer Engineering, Faculty of Applied Mathematics and Computer Science, Faculty of Humanities it is enough to carry out the classification with sufficient accuracy. For other faculties such attributes as "Social welfare", "A category of secondary school", "Place of residence" and "An average USE score on admission" are displayed from the second to the fifth level of the tree, which means that these attributes affect most of the classifications of the instances, but this effect is much less than that of the attribute "Average score in the first session". At the same time, the attribute "Enrollment" does not appear in the classification tree for any faculty or any faculty program, which indicates that it does not affect the results of the classification.

The following conclusions can be made from the research results in general:

- The considered input characteristics have different effect on the dropout of students. The decisive factor is "Average score in the first session" among those factors considered influencing the results of the classification. Other attributes, including "An average USE score on admission," do not have a significant effect on student dropout.

- In most cases the absence of the attribute "Category of secondary school" in constructed decision trees indicates that this attribute has a rather weak effect on student performance at the university.

Table 1: CLASSIFICATION RESULTS FOR EDUCATIONAL PROGRAM

Support Vector Machine algorithm				
Faculties / Educational programs	Sample size	Variant1	Variant2	
		Accuracy	Accuracy on the training sample	Accuracy on the test sample
Faculty of Automation and Computer Engineering	1178	78.9	79.5	79.6
09.03.01	393	79.6	80.1	78.6
09.03.02	125	84.8	85.5	69.0
09.03.04	98	81.6	83.8	83.1
.....	.....	.....	.....	.....
Faculty of Applied Mathematics and Computer Science	641	81.1	82.4	77.1
01.03.02	447	80.8	82.2	75.2
02.03.03	194	81.6	85.3	73.8
Aircraft Faculty	1001	78.8	79.7	76.3
05.03.06	33	93.9	95.4	63.6
15.03.03	109	85.3	88.1	78.4
.....	.....	.....	.....	.....
Faculties/Educational programs	CART algorithm		Random Forest algorithm	
	Accuracy on the training sample	Accuracy on the test sample	Accuracy on the training sample	Accuracy on the test sample
Faculty of Automation and Computer Engineering	79.1	78.1	77.9	80.4
09.03.01	78.6	77.9	77.1	75.4
09.03.02	83.1	73.8	81.9	78.5
09.03.04	55.3	45.5	80.0	72.7
.....	.....	.....	.....	.....
Faculty of Applied Mathematics and Computer Science	79.1	78.1	77.9	80.4
01.03.02	78.5	76.5	78.5	75.5
02.03.03	81.4	77.9	81.5	80.0
Aircraft Faculty	78.5	77.5	80.8	76.4
05.03.06	68.2	45.5	77.3	65.5
15.03.03	58.3	54.1	86.6	78.4
.....	.....	.....	.....	.....

- The low value of the attribute "Average score in the first session" indicates a potential threat of future student dropout with a high probability. Such students with a high threat of future dropout should be provided with additional assistance measures (special conditions of study, additional consultations, etc.) at the early stage (after the first session) in order to retain them at the university.

In addition, the dynamics of changes in the significance of the studied factors in the time range and the level of influence of the transition from numerical to categorical data were investigated.

The accuracy of prediction could be increased by considering additional input factors: the fact of the student's work in parallel with the studying, the student's personal qualities (motivation, insistence, etc.), school academic performance related to the student's profile. Unfortunately, these data are not currently available in the Information System. In addition, factors such as the social status of the student's family, qualifications, education and family income of parents would be useful, but the collection of this data is impossible under the Federal Law on Personal Data.

3 Analysis of the demand for educational resources in the Electronic library system of NSTU

Creation of electronic educational resources (manuals, lecture notes, etc.) is one of the activities of the teaching staff of the university. Electronic library system of NSTU contains several tens of thousands of resources written by scientific and pedagogical staff of the university. However, the level and quality of these resources very vary. Just as the level of scientific publications are estimated by the citation index, it is logical to assess the quality of educational resources by their demand. The demand of educational resources estimated by the number of downloads these resources by students.

This allows you to get the ratings of the most popular electronic educational resources and ratings of authors whose resources are most in demand. The qualitative dependences of these ratings are shown in Fig. 3 and Fig. 4.

At NSTU, as at most universities, the system of effective contract introduced and over the past few years used. Two factors dominated with the highest weights at the system: they are the publication activity of employees and the amount of money earned for the university. Twice a year, the results of executing of the effective contract faculties, departments and formed a rating of effective contract research and teaching staff of the university publicly summer up. Herewith the administration of the University constantly notes that about one hundred of the most active employees of the university perform 50% of the indicators of an effective contract. This indicates, according to the administration, the inefficient work of a large part of the scientific and pedagogical staff of the university. In this regard, it is interesting to correlate the performance indicators of the effective contract of scientific and pedagogical workers and the demand for electronic educational resources of these employees. At Fig. 5 data on the number of downloads of the authors' resources were superimposed on data

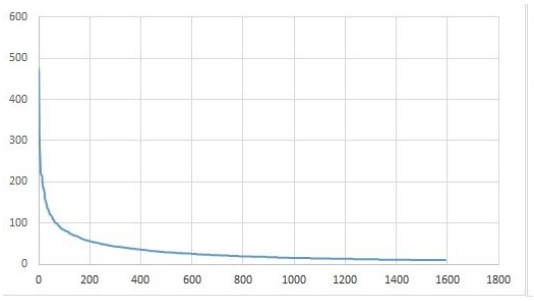


Figure 3: Download resources of Electronic library system of NSTU (the most popular resources)

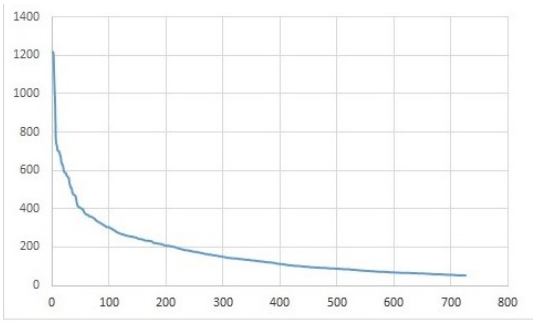


Figure 4: Download resources of Electronic library system of NSTU (the most popular authors)

of scientific and pedagogical staff' rating. We see that the demand for educational resources of scientific and pedagogical staff from the first thousand rating keep at about the same level, after that they sharply reduce.

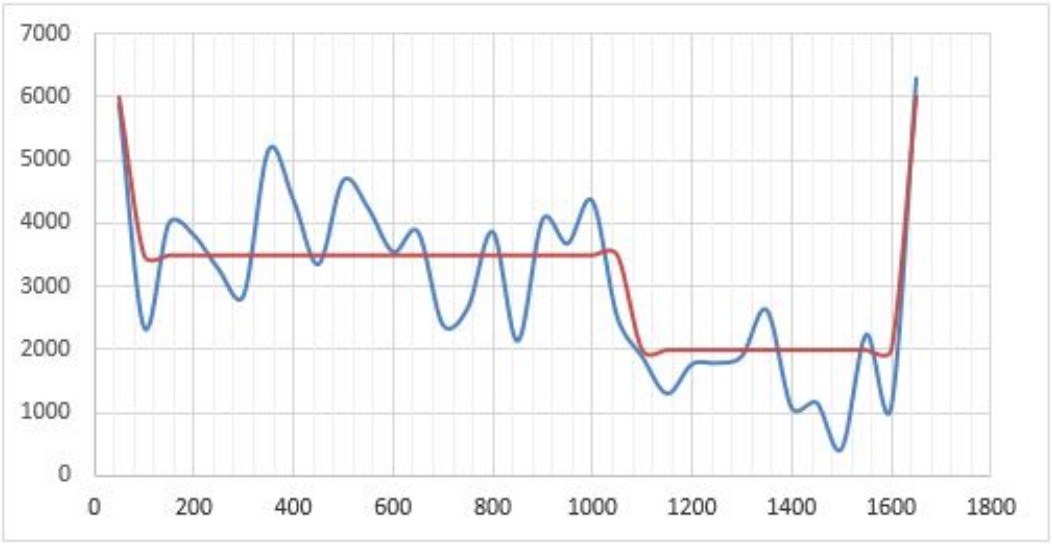


Figure 5: Correlation of execution of effective contract and the demand for electronic educational resources of scientific and pedagogical workers

The higher demand for educational resources of employees who lead the rating of an effective contract can explained by the fact that their training materials are prepared in collaboration with other employees. Herewith a demand for educational resources of those employees who close the rating are increase dramatically. This indicates that not succeed in scientific and research activities, these employees are very useful for the university in the preparation of high-quality electronic educational resources.

4 Research of approaches to improving the quality of higher education, carried out using Methods of Expert Estimations

One of directions of research is the analysis of approaches to improving the quality of higher education, carried out using Methods of Expert Estimations.

As a result, of preliminary conversations with university professors with extensive teaching experience, eight ways were formulated to solve this problem:

1) To eliminate the direct relationship between the number of university students and the amount of funding from the ministry, which would allow the university to get rid of careless students

2) To involve in the process of teaching in a higher educational institution highly qualified specialists from production, research and commercial organizations on conditions of adequate remuneration.

3) To increase the level of knowledge in the admission of applicants, which will allow to cut off poorly trained high school graduates in the entrance examinations.

4) To increase the level of salary for university teacher.

5) To improve the qualifications of university professors through active use of the internship system, retraining, and improving scientific qualifications.

6) To enroll more students to participate in scientific, design and engineering works of industrial or scientific importance.

7) To refuse to blindly copy the Western-style higher education system.

8) To improve the quality of study courses and their methodological support by prior obligatory discussion of the courses developed by the teacher within the pedagogical community of the university.

The study's expert group consisted of 13 employees of the NSTU with work experiment between 7 to 48 years:

The experts represented

- three faculties: the Faculty of Applied Mathematics and Computer Science, the Faculty of Mechanical engineering and Technologies and the Faculty of Mechatronics and Automation;

- four departments: Department of Applied Mathematics, Department of Theoretical and Applied Computer Science, Department of Materials Science in Mechanical Engineering, Department of theoretical fundamentals of electrical engineering.

There were three heads of departments, five professors, six associate professors, one senior lecturer and three heads of university departments among the study's experts.

All experts ranked from 3.7 to 10.8, depending on the work experience service, position, academic degree. The experts' preferences were processed by five methods of expert assessments after conducting the survey: the method of pairwise comparisons, the method of weighted expert assessments, the method of preference, the rank method, the method of complete pairwise comparison.

Among the most unexpected results stands the fact that all five methods of expert assessments gave preference to the alternative "Increasing the level of university teacher's salary", despite the fact that among the proposed alternatives were alternatives that deserve more attention. At the same time, it cannot said that mercantile interest won, since the most of the experts did not put this alternative in the first place.

It addition it should be noted that various experts expressed very different preferences, which indicates the absence of a unified university strategy in approaches to solving the problem of improving the quality of higher education.

References

[1] Kabakchieva D. (2013). Predicting student performance by using data mining methods for classification. *Cybernetics and Information Technologies.*. Vol. **13**, pp. 61-72.

[2] Kabakchieva D., Stefanova K., Kisimov V. (2012). Analyzing University Data for Determining Student Profiles and Predicting Performance. *Conference Proceedings of the 4th International Conference on Educational Data Mining (EDM 2011)*. pp. 347-348.

[3] Pal S. (2012).Mining educational data to reduce dropout rate of engineering students. *International engineering electronic business*. No. **2**, pp. 1-7.

[4] Stasyshin V.M., Avrunev O.E., Afonina E.V., Lyax K.N. (2012). Informacionnaja sistema universiteta: opyt sozdanija i tekushhee sostojanie [Information system of the University: experience and current status]. *Otkrytoe i distancionnoe obrazovanie [Open and distance education]* . pp. 9-15.

[5] Stasyshin V.M., Stasyshin T.V. (2018) Analysis of educational data in the decision-making support system of university. *Proceedings of the 14th International Conference on Actual problems of electronic instrument engineering (APEIE–2018)*. Vol. **1**, No. **4**, pp. 541-545.

[6] Superby J. Vandamme J., Meskens N. (2006) Determination of factors influencing the achievement of the first-year university students using data mining methods.*Proceedings of the Workshop on Educational Data Mining at the 8th International Conference on Intelligent Tutoring Systems (ITS 2006)*. pp. 37-44.

Adaptive Kernel identification of nonlinear stochastic dynamical systems

Antropov N. and Agafonov E.
Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russia
e-mail: nikita.antropov.92@mail.ru, agafonov@gmx.de

Abstract

The paper considers the identification of nonlinear stochastic dynamical systems. One approach for adaptive identification of nonlinear dynamical systems based on kernel adaptive filtering is proposed. Simulations illustrating performance for one stochastic dynamical system are presented. Convergence rate and prediction accuracy are analyzed for various noise levels in data.

Keywords: identification, kernel, stochastic system, dynamical system

Introduction

Nowadays one key problem is identification of nonlinear dynamical systems with stochastic behavior. Most of up-to-date identification approaches are based on linear models since their properties and limits are well-known and established. At the same time, quite often one has to deal with nonlinear systems that cannot be identified by linear methods. In this regard, kernel methods for identification of nonlinear systems have become widely used since they allow using linear algorithms to solve nonlinear identification problems without any assumptions about the model structure.

Over the past decades there have been developed many kernel-based algorithms, which have become widely used in identification of dynamical systems. Most of researches in the field of kernel adaptive identification present results of simulations illustrating the rate of convergence, accuracy and comparison of the amount of calculations and memory requirements. In many cases, computational experiments are performed for not optimized parameters that may cause high bias of obtained results. The purpose of this paper is to analyze properties of adaptive kernel algorithms at optimal parameter settings on identification problem of stochastic dynamical systems.

1 Kernel methods

Kernel methods rely on the so-called kernel trick [4]. The key feature of kernalization is that complex and nonlinear functions in original input space $\mathcal{X}$ are more likely to be linear in a high dimensional Hilbert space $\mathcal{H}$, also called feature space. The original input space $\mathcal{X}$ can be mapped to feature space $\mathcal{H}$ by using mapping $\phi : \mathcal{X} \to \mathcal{H}$. The mapping ϕ can be defined explicitly, if a prior knowledge is available. In temrs of kernel methods, the mapping is defined implicitly from the training data by a positively defined kernel function $\mathcal{K}(\mathbf{x}_1, \mathbf{x}_2) = \phi(\mathbf{x}_1)^{\mathrm{T}}\phi(\mathbf{x}_2)$ satisfying Mercer's conditions [4].

This transformation allows solving nonlinear problems through construction of kernelized counterparts of linear methods by simply replacing inner products in linear algorithm with kernel.

The core of kernel methods is Representer Theorem [2], that allows solving nonlinear optimization problems through construction a linear combination of kernel function on training data:

$$\hat{y}(\mathbf{x}) = \sum_{n=1}^{N} \alpha_n \mathcal{K}(x_n, \mathbf{x}), \tag{1}$$

which parameters α_n can be found by methods based on least squares or least-mean squares criterion, like kernel ridge regression [3] and kernel least-mean squares [4]. Next section presents a briefly review of such algorithms in notations proposed in [5].

1.1 Kernel ridge regression

Assume $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_n]^{\mathrm{T}} \in \mathbb{R}^{n \times m}$ as input matrix, $\mathbf{y} = [y_1, ..., y_n]^{\mathrm{T}} \in \mathbb{R}^{n \times 1}$ as output vector, where n is a number of training data, m is a number of inputs. Regularized least squares problem, or ridge regression, involves seeking a weight vector $\mathbf{w} = [w_1, ..., w_m] \in \mathbb{R}^{1 \times m}$ that solves:

$$\min_{\mathbf{w}} J = \min_{\mathbf{w}} \|\mathbf{y} - \mathbf{X}\mathbf{w}\|^2 + c\|\mathbf{w}\|^2, \tag{2}$$

where c is a positive Tikhonov regularization constant. Solution of the problem (2) for $\mathbf{w}$ is:

$$\mathbf{w} = (\mathbf{X}^{\mathrm{T}}\mathbf{X} + c\mathbf{I})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}, \tag{3}$$

where $\mathbf{I}$ is identity matrix.

Regularization improves stability of matrix $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ allowing the solution of ill-posed problems and avoids overfitting by trying to keep norm of the vector $\mathbf{w}$ small.

In order to construct a kernel-based representation of kernel ridge regression, transform matrix $\mathbf{X}$ and vector $\mathbf{w}$ in equation (2) into the feature space:

$$\min_{\phi(\mathbf{w})} J = \min_{\phi(\mathbf{w})} \|\mathbf{y} - \phi(\mathbf{X})\phi(\mathbf{w})\|^2 + c\|\phi(\mathbf{w})\|^2. \tag{4}$$

According to Representer Theorem [2], the solution $\phi(\mathbf{w})$ of this problem can be expressed as linear combination of training data in feature space:

$$\phi(\mathbf{w}) = \phi(\mathbf{X}^{\mathrm{T}})\boldsymbol{\alpha}, \tag{5}$$

where $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_n]^{\mathrm{T}}$. Then substitute (5) in (4), denote $\mathbf{K} = \phi(\mathbf{X}^{\mathrm{T}})\phi(\mathbf{X})$, and obtain the kernel-based least squares problem:

$$\min_{\boldsymbol{\alpha}} J = \min_{\boldsymbol{\alpha}} \|\mathbf{y} - \mathbf{K}\boldsymbol{\alpha}\|^2 + c\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{K}\boldsymbol{\alpha}, \tag{6}$$

where vector $\boldsymbol{\alpha}$ can be found by:

$$\boldsymbol{\alpha} = (\mathbf{K} + c\mathbf{I})^{-1}\mathbf{y}. \tag{7}$$

Prediction of the output $\hat{y}_n$ at time n is given by:

$$\hat{y}_n = \mathbf{k}_n^{\mathrm{T}} \boldsymbol{\alpha}, \tag{8}$$

where $\mathbf{k}_{\mathrm{n}}^{\mathrm{T}} = [\mathcal{K}(\mathbf{x}_n, \mathbf{x}_1), ..., \mathcal{K}(\mathbf{x}_n, \mathbf{x}_{n-1})]$.

Predictions obtained by kernel ridge regression algorithm requires $O(n^3)$ calculations, when its recursive version, kernel recursive least squares (KRLS) algorithm [6], requires about $O(n^2)$ calculations, which is more preferred in online scenarios. Consideration of KRLS algorithm is presented in the next section.

1.2 Kernel recursive least squares

In order to construct a KRLS algorithm [6], assume that data arrive sequentially, one pair $\mathbf{x}_n, y_n$ at time n. Then, vector $\boldsymbol{\alpha}$ and prediction $\hat{y}_n$ at time n can be written as:

$$\boldsymbol{\alpha}_{n-1} = \mathbf{K}_{r,n}^{-1} y_{n-1}, \quad \hat{y}_n = \mathbf{k}_n^{\mathrm{T}} \boldsymbol{\alpha}_{n-1} \tag{9}$$

where $\mathbf{K}_{r,n} = \mathbf{K}_n + c\mathbf{I}$.

By using prediction $\hat{y}_n$ and obtained output y_n, a priori error e_n at time n is calculated:

$$e_n = y_n - \hat{y}_n. \tag{10}$$

Then kernel matrix $\mathbf{K}_{r,n}$ is updated:

$$\mathbf{K}_{r,n} = \begin{bmatrix} \mathbf{K}_{r,n-1} & \mathbf{k}_n \\ \mathbf{k}_n^{\mathrm{T}} & \mathbf{k}_{nn} + c \end{bmatrix}, \tag{11}$$

where $\mathbf{k}_{nn} = \mathcal{K}(\mathbf{x}_n, \mathbf{x}_n)$.

By introducing the variables:

$$\boldsymbol{\beta}_n = \mathbf{K}_{r,n-1}^{-1} \mathbf{k}_n, \quad \boldsymbol{\gamma}_n = \mathbf{k}_{nn} + c - \mathbf{k}_n^{\mathrm{T}} \boldsymbol{\beta}_n. \tag{12}$$

the inverse kernel matrix is updated:

$$\mathbf{K}_{r,n}^{-1} = \frac{1}{\boldsymbol{\gamma}_n} \begin{bmatrix} \boldsymbol{\gamma}_n \mathbf{K}_{r,n-1}^{-1} + \boldsymbol{\beta}_n \boldsymbol{\beta}_n^{\mathrm{T}} & -\boldsymbol{\beta}_n \\ -\boldsymbol{\beta}_n & 1 \end{bmatrix}. \tag{13}$$

Finally, KRLS prediction $\hat{y}_n$ and update of the vector $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_n]^{\mathrm{T}}$ is performed by the recursive algorithm:

$$\hat{y}_n = \mathbf{k}_n^{\mathrm{T}} \boldsymbol{\alpha}_{n-1}, \tag{14}$$

$$\boldsymbol{\alpha}_n = \begin{bmatrix} \boldsymbol{\alpha}_{n-1} \\ 0 \end{bmatrix} - \frac{e_n}{\boldsymbol{\gamma}_n} \begin{bmatrix} \boldsymbol{\beta}_n \\ -1 \end{bmatrix}. \tag{15}$$

1.3 Kernel least-mean squares

Consider a least squares problem in feature space (3). Its solution $\mathbf{w}$ can be obtained by using stochastic-gradient method [7]:

$$\phi(\mathbf{w}) \leftarrow \phi(\mathbf{w}) - \frac{\eta}{2}\frac{\partial J}{\partial \phi(\mathbf{w})}, \tag{16}$$

where η is a learning rate.

After replacing derivate by its instantaneous approximation and removing regularization, an update rule for kernel least-mean squares (KLMS) [4] is obtained:

$$\phi(\mathbf{w}_n) = \phi(\mathbf{w}_{n-1}) + \eta e_n \phi(\mathbf{x}_n). \tag{17}$$

Finally, KLMS prediction $\hat{y}_n$ and update of the vector $\boldsymbol{\alpha}$ is calculated by:

$$\hat{y}_n = \mathbf{k}_n^{\mathrm{T}} \boldsymbol{\alpha}_{n-1}, \quad \boldsymbol{\alpha}_n = \begin{bmatrix} \boldsymbol{\alpha}_{n-1} \\ \eta e_n \end{bmatrix}. \tag{18}$$

If there is a regularization, we obtain so called naive online regularized risk minimization algorithm (NORMA), which prediction $\hat{y}_n$ and update rule for $\boldsymbol{\alpha}$ is given by:

$$\hat{y}_n = \mathbf{k}_n^{\mathrm{T}} \boldsymbol{\alpha}_{n-1}, \quad \boldsymbol{\alpha}_n = \begin{bmatrix} (1-\eta c)\boldsymbol{\alpha}_{n-1} \\ \eta e_n \end{bmatrix}. \tag{19}$$

KLMS and NORMA requires only $O(n)$ calculations, what can be more appropriate in low cost systems.

2 Problem statement and simulation setup

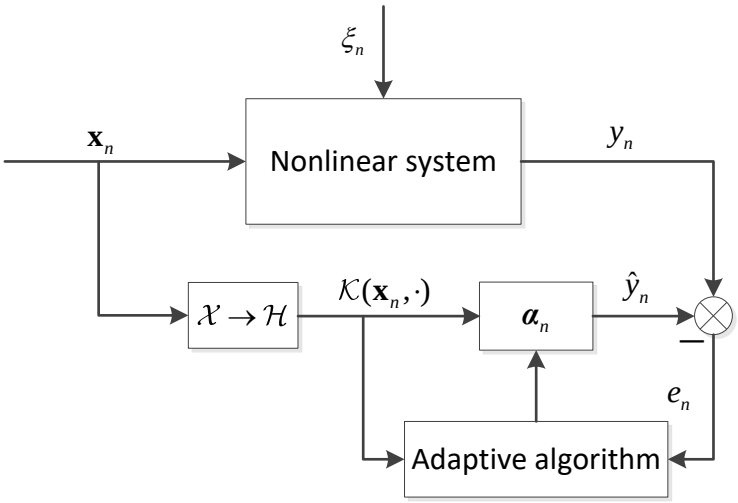


Figure 1: Kernel-based method identification setup [7].

Directly applying kernel-based approach to the identification problem, the next problem statement is obtained. Consider a nonlinear system $y_n = f(\mathbf{x}_n) + \xi_n$ influenced by unobservable random effects ξ_n having zero mean and limited variance, and a set of observable $\mathbf{x}_n$. Output of the system y_n is synchronically measured together with observable input $\mathbf{x}_n$. Input $\mathbf{x}_n$ and output y_n arrive sequentially, one data instance at time $n = 1, 2, ...$ Typical setup for online system identification with a kernel-based method is shown in Fig. 1.

Every identification algorithm have a number of parameters to be chosen properly in order to obtain lower bias. In general, parameters should be optimized within selected criterion. There are many approaches to parameter optimization, including typical cross-validation, marginal likelihood maximization, adaptive optimization using stochastic-descent method and etc.

Parameters influence on prediction accuracy and convergence speed is one key problem of the proposed adaptive kernel identification algorithms, which is necessary to specify. Further simulation presents performance results for NORMA and KRLS algorithms at optimal settings of parameters on identification problem of one nonlinear dynamical system affected by additive noise.

Simulations are performed for a Mackey-Glass time series [9], generated by the following equation:

$$\frac{dx_n}{dn} = -bx_n + \frac{ax_{n-\Delta}}{1 + x_{n-\Delta}^{10}} + \xi_n, \tag{20}$$

where $a = 0.2$,$b = 0.1$, $\Delta = 30$, ξ_n is an additive noise with mean $E[\xi_n] = 0$ and variance $\sigma = 0.1$.

Generation of training and test samples is performed by 4-th order Runge-Kutte method for initial state $x(0) = 1.2$ and step size $h = 1$.

Sample size $N = 1250$, first 1000 observations are used for training and last 250 for testing, noise levels $\xi_n = 0\%$, 5%, 10%, 15%.

Consider a time-delayed vector $\mathbf{x}_n = [x_{n-1}, x_{n-2}, x_{n-3}]^{\mathrm{T}}$ as input data, $y_n = x_n$, $n = 1, 2, ..., N$ as output data.

Assume Gaussian kernel as a kernel function:

$$\mathcal{K}(\mathbf{x}_1, \mathbf{x}_2) = \exp\left(-\frac{\|\mathbf{x}_1 - \mathbf{x}_2\|^2}{2\sigma^2}\right). \tag{21}$$

Simulations are performed for optimized lenght-scale parameter σ, regularization parameter c and NORMAs step size η obtained by cross-validation performed offline in terms of normalized mean-squared error (nMSE):

$$\mathrm{nMSE} = \log_{10}\left[\frac{1}{N\hat{\sigma}^2}\sum_{n=1}^{N}(y_n - \hat{y}_n)^2\right]. \tag{22}$$

In next section simulations results, including learning curves and optimal parameters configuration at various noise levels are presented.

3 Simulations results

Next figures illustrate convergence curves for KRLS and NORMA obtained on the test sample throughout the training run.

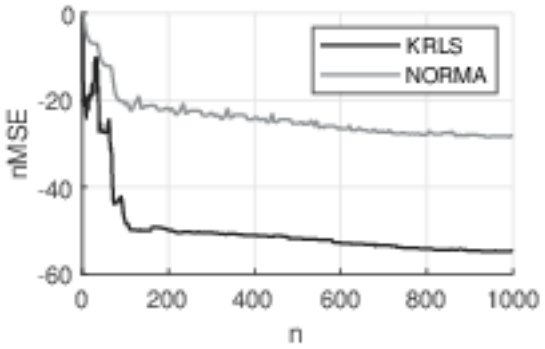


Figure 2: Learning curves at $\xi_n = 0\%$.

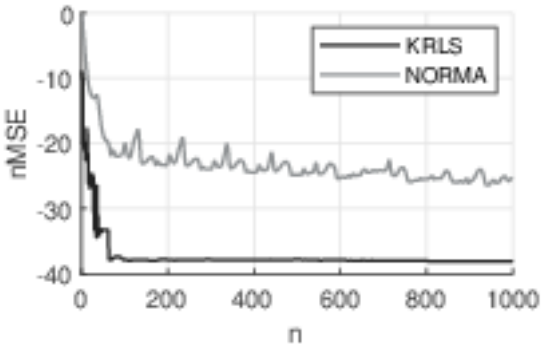


Figure 3: Learning curves at $\xi_n = 5\%$.

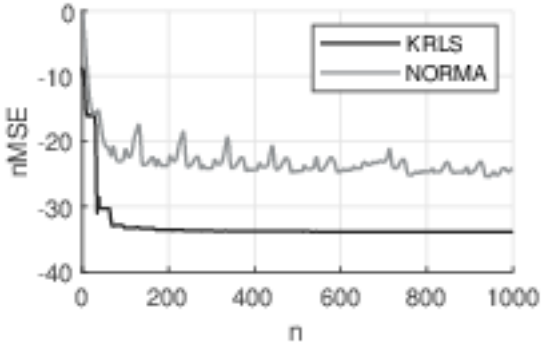


Figure 4: Learning curves at $\xi_n = 10\%$.

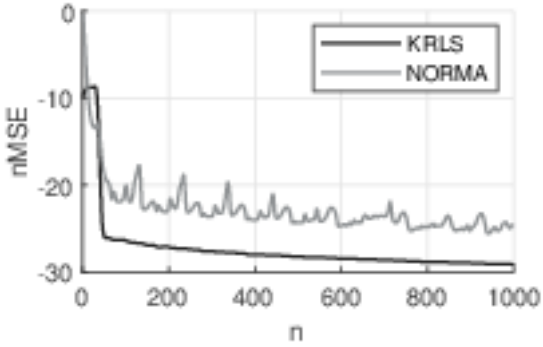


Figure 5: Learning curves at $\xi_n = 15\%$.

Figures 2-5 shows that KRLS has a higher convergence speed and lower steady-state nMSE on every noise level, that caused by its higher complexity. KRLS updates all parameters $\boldsymbol{\alpha}$ and inverse kernel matrix $\mathbf{K}_{r,n}^{-1}$ at each iteration, while NORMA updates only one parameter α_n.

Raising regularization parameter c for KRLS in table 1 indicates that with an increasing noise level the solution becomes more unstable. In case of NORMA optimal regularization parameter c is unchanged, because KLMS-type algorithms have a "self-regularization" property, when their parameter η is small enough. In order to provide smoothing, lenght-scale parameter σ also increases on each noise level for both algorithms. Optimal length-scale parameters σ for NORMA is lower than optimal σ for KRLS. It shows that KRLS take into account more data from training set at prediction point than NORMA, but it is still able to obtain lower nMSE.

Figure 6 presents predictions obtained by KRLS and NORMA on test data at $\xi_n = 15\%$. It can be seemed that NORMA has a tendency to oversmooth data where peaks are observed, while KRLS prediction is more consistent with real output.

Table 1: Optimized parameters and nMSE per noise level

	NORMA				KRLS		
	σ	c	η	nMSE	σ	c	nMSE
$\xi_n = 0\%$	0.5749	$8.17 \cdot 10^{-12}$	0.0817	-27.95	4.0456	$5.64 \cdot 10^{-20}$	-40.51
$\xi_n = 5\%$	0.5979	$8.17 \cdot 10^{-12}$	0.0778	-25.64	5.8444	$8.54 \cdot 10^{-7}$	-28.14
$\xi_n = 10\%$	0.6547	$8.17 \cdot 10^{-12}$	0.0709	-22.22	6.9932	$4.21 \cdot 10^{-5}$	-23.68
$\xi_n = 15\%$	0.8812	$8.17 \cdot 10^{-12}$	0.0642	-19.97	7.9412	0.0058	-20.45

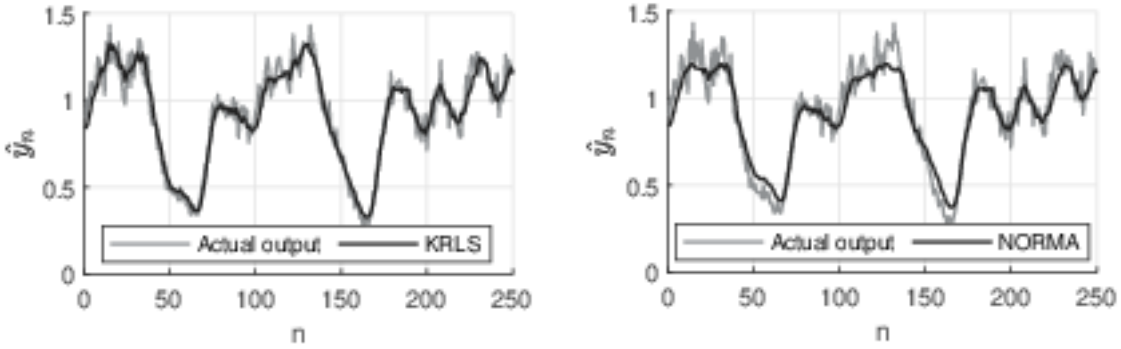


Figure 6: KRLS and NORMA prediction on test data at $\xi_n = 15\%$.

One may noticed that NORMA has fluctuations in nMSE, that is caused by constant step-size η, which should be chosen properly to yield convergence of the stochastic-gradient method. NORMA rely on least-mean criterion, thus it is more persistent to noise effects and has lower nMSE spread.

Conclusions

Kernel methods are nonparametric techniques that allow solving nonlinear identification problems without any model specification. In this paper the adaptive kernel-based identification problem of nonlinear dynamical systems is considered, including its two basic methods: kernel recursive least squares and kernel least-mean squares. The proposed algorithms were compared during the numerical simulations at optimal parameter configurations. In the simulations it is shown that the strength of kernel least-mean squares algorithm lies in its ability to persist noise effects upon sufficiently low prediction error in addition to its computational efficiency, while kernel recursive least squares algorithm has lower prediction error, but it is more computationally complex. Selection of practical identification algorithm is a trade-off between accuracy and computational complexity. Obtained results give reason to design modi-

fications and to perform application of kernel adaptive identification algorithms. This paper affords a basis for further research and more detailed analysis of kernel adaptive identification algorithms.

References

[1] Aizerman M.A., Braverman E. M., Rozoner L. (1964). Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control*. Vol. **25**, pp. 821-837.

[2] Schölkopf B., Herbrich R., Smola A.J. (2001). A generalized representer theorem. *Computational learning theory*. Vol. **2111**, pp. 416-426.

[3] Saunders C., Gammerman A., Vovk V. (1998). Ridge regression learning algorithm in dual variables. *Proceedings of the 15th International Conference on Machine Learning (ICML)*. pp. 515-521.

[4] Liu W., Pokharel P.P., Principe J.C. (2008). The kernel least-mean-square algorithm. *IEEE Transactions on Signal Processing*. Vol. **56**, No. **2**, pp. 543-554.

[5] Suykens J.A.K., Signoretto M., Argyriou A.(2014). *Regularization, Optimization, Kernels, and Support Vector Machines*. Chapman & Hall/CRC, London.

[6] Engel Y., Mannor S., Meir R. (2004). The kernel recursive least squares algorithm. *IEEE Transactions on Signal Processing*. Vol. **52**, No. **8**, pp. 2275-2285.

[7] Haykin S. (2002). *Adaptive Filter Theory*. NJ: Prentice-Hall, Englewood Cliffs.

[8] Parreira W.D., Bermudez J.-C.M., Richard C., Tourneret J.-Y. (2012). Stochastic behavior analysis of the Gaussian kernel least-mean-square algorithm. *IEEE Transactions on Signal Processing*. Vol. **60**, No. **5**, pp. 2208–2222.

[9] Liu W., Principe J.C., Haykin S. (2010). *Kernel Adaptive Filtering: A Comprehensive Introduction*. JohnWiley & Sons.

An optimal design of the experiment in the active identification of locally adaptive linear regression models

Alexander A. Popov and Victoria M. Volkova
Novosibirsk state technical university, Novosibirsk, Russia
e-mail: a.popov@corp.nstu.ru, volkova@ami.nstu.ru

Abstract

To restore dependencies, it is proposed to use locally adaptive linear regression models. The question of a priori optimal design of the experiment when identifying locally adaptive linear regression models is considered.

At the same time, domains of the the operating factors is divided into 2-3 fuzzy partitions. The problem of construction D-optimal design of the experiment is formulated. The problem of construction D-optimal design is considered for the case of one and two factors with the number of fuzzy partitions 2 and 3. The analysis of the optimal designs characteristics depending on the fuzzy partitions intersection zone width is carried out. It is noted that with a decrease in the fuzzy partitions intersection zone, efficiency of optimal designs increases, which affects the reduction of the dispersion matrices determinants. The conclusion is made about the efficiency of active identification of locally adaptive linear regression models.

Keywords: locally adaptive linear regression models, membership function, optimal design of experiment, criterion of D-optimality.

Introduction

In real situations, knowledge about the object model at the initial stage of the study is far from complete. In the conditions of complete or partial ignorance of the structure of an object model, in practice often used the methodology of modeling and experimentation on the principle from "simple" to "complex", for example, from linear to quadratic model, etc. In accordance with the stages of complication of the model and the experimental design is selected.

If the range of the factors variation is chosen wide enough, then one should expect that the response behavior in different parts of the factor space will differ. In conditions when a priori there are no unambiguous assumptions of structure of object model, they often resort to using non-parametric regression modeling technology, when selection of model is carried out according to the available data. As an example of such technology it is possible to consider a support vector machine (see, for example, [1]- [3]). However, there is a rather effective methodology, which, while remaining within the framework of a linear parametric model, can obtain descriptions of the behavior of an object on rather broad domain of actions. It is a technology of locally adaptive regression models, when the behavior of an object in different parts of the factor space is modeled by separate local models [4]- [9]. Their efficiency

can be also increased due to use of optimal experimental designs for such models. However, the theory of optimal experimental design for such models has not yet been developed.

1 Materials and methods

Basic assumptions of the theory of optimal experimental design. Let's assume that the studied observation model is

$$y = f^T(x)\theta + e = \sum_{l=1}^{m} f_l(x)\theta_l + e, \tag{1}$$

where $f^T(x) = (f_1(x), ..., f_m(x))-$ vector of known functions from an explanatory variable $x = (x_1, ..., x_k)^T$, which can change in range $\tilde{X}$; $\theta = (\theta_1, ..., \theta_m)^T$ - unknown parameters; e - error term; y - value of an dependent variable. Let results of measurements be independent random variables with the expected value determined by the regression equation (response function)

$$E(y/x_j) = f^T(x_j)\theta = \eta(x_j, \theta) \tag{2}$$

and dispersion σ_j^2 in each point $x_j \in \tilde{X}$, $j = \overline{1, N}$. For the errors e_j at $x_j \in \tilde{X}$ therefore it is possible to write $E(e_j) = 0$; $E(e_j e_k) = \sigma_j^2 \delta_{jk}$, where $\delta_{ik}-$ Kronecker delta, $j,\ k = \overline{1, N}$.

The priori choice of points x_j, $j = \overline{1, N}$, according to these or those optimality criterions also there is a problem of optimal experimental design.

Under the discrete normalized design ε_N we will understand the set of quantities $x_1, x_2, ..., x_n$; $p_1, p_2, ..., p_n$, where $\sum\limits_{i=1}^{n} p_i = 1$, $p_i = r_i/N$:

$$\varepsilon_N = \left\{ \begin{matrix} x_1, x_2, ..., x_n \\ p_1, p_2, ..., p_n \end{matrix} \right\}, \quad \sum_{i=1}^{n} p_i = 1, \quad p_i = r_i/N, \quad i = \overline{1, n}.$$

If the weights of the design points are imposed only restrictions $\sum\limits_{i=1}^{n} p_i = 1$, $p_i \geq 0$, $i = \overline{1, n}$, then such a design is called a continuous normalized design ε. Further, without loss of generality, we will assume that the observations are uniformally precise, i.e. $\sigma_i^2 = \sigma_j^2$, $\forall i, j = 1, \ldots n$. For the design ε information matrix takes the form

$$M(\varepsilon) = \sum_{i=1}^{n} p_i f(x_i) f^T(x_i). \tag{3}$$

The experimental design ε (or ε_N), chosen in a certain way, allows, in accordance with (3), to improve parameter θ estimation accuracy.

We will evaluate the quality of the design ε by the value of a certain functional of the information matrix $M(\varepsilon)$ or the corresponding to it dispersion matrix $D(\varepsilon) =$

$M^{-1}(\varepsilon)$. In this work we will consider a task of constructing D-optimal designs: $\varepsilon^* = Arg \max_{\varepsilon} |M(\varepsilon)|$.

The necessary and sufficient condition for the D-optimality of the design ε^* is to fulfill [10]:

$$\max_{x \in \tilde{X}} \varphi(x, \varepsilon^*) = trM(\varepsilon^*) \frac{\partial \ln |M(\varepsilon^*)|}{\partial M(\varepsilon)} = trM(\varepsilon^*) M^{-1}(\varepsilon^*) = m, \tag{4}$$

where $\varphi(x, \varepsilon^*) = f^T(x) M^{-1}(\varepsilon^*) f(x)$.

The construction of optimal experimental designs is carried out, as a rule, using the appropriate numerical procedures. In this case, the execution of of conditions (4) is achieved with some permissible accuracy:

$$\left| - \min_{x \in X} \varphi(x, \varepsilon^s) + trM(\varepsilon^s) \frac{\partial \psi[M(\varepsilon^s)]}{\partial M(\varepsilon^s)} \right| \leq \delta, \tag{5}$$

where δ - small positive quantity.

In this paper, to obtain optimal designs, we will use the gradient projection method based on the weights of the design points. The spectrum of the initial design itself will be a discrete set of points in the form of a sufficiently dense grid on $\tilde{X}$. We will control the achievement of the extremum point by execution of ratio (5).

Locally adaptive regression models. Locally adaptive regression models (LAR models) will be specified through the regression tree. Let $x_1,\ x_2,\ \ldots, x_k$ represent linguistic variables. Their values are determined by fuzzy sets $A,\ B, \ldots, \Gamma$, and the degree of intensity of the manifestation of the value will be set through membership functions. The branches of the decision tree have the form [11]

$$\begin{array}{l} \Pi_{ij..l}: \ If\ (x_1\ is\ A_i) \wedge (x_2\ is\ B_j)\ \wedge ... \wedge (x_k\ is\ \Gamma_l)\ \text{then} \\ y'_{ij...l} = \eta + \alpha_i + \beta_j + ... + \gamma_l. \end{array} \tag{6}$$

Truth of statements $(x_1\ is\ A_i),\ (x_2\ is\ B_j),\ ....,(x_k\ is\ \Gamma_l)$ is defined by values of the corresponding membership functions $\mu_{A_i} \in [0,\ 1],\ \mu_{B_j} \in [0,\ 1], \ldots, \mu_{\Gamma_l} \in [0,\ 1]$. Degree of truth of the statement $\Pi_{ij...l}$ will be denoted as $\mu(y'_{ij..l})$ and calculate it as $\mu(y'_{ij..l}) = \mu_{A_i} \mu_{B_j} \cdots \mu_{\Gamma_l}$. Concerning assignment of values $\mu_{A_i}, \mu_{B_j}, \ldots, \mu_{\Gamma_l}$ we make the requirement that for each observation conditions are satisfied:

$$\begin{array}{l} \sum\limits_{i=1}^{I} \mu_{A_i} = 1,\ \sum\limits_{j=1}^{J} \mu_{B_j} = 1,\ ...\ ,\ \sum\limits_{l=1}^{L} \mu_{\Gamma_l} = 1; \\ \mu_{A_i} \in [0,1],\ i = \overline{1,I},\ \mu_{B_j} \in [0,1],\ j = \overline{1,J},\ \ \ldots,\ \mu_{\Gamma_l} \in [0,1],\ l = \overline{1,L}. \end{array} \tag{7}$$

Taking into account (7) tree decisions (6) can be represented as an observation model

$$y_{ij...l} = \eta + \sum_{i=1}^{I} \mu_{A_i} \alpha_i + \sum_{j=1}^{J} \mu_{B_j} \beta_j + ... + \sum_{l=1}^{L} \mu_{\Gamma_l} \gamma_l + e_{ij...l}. \tag{8}$$

After finding the estimate $\hat{\theta}^T = (\hat{\eta},\ \hat{\alpha}_1, ..., \hat{\alpha}_I,\ \hat{\beta}_1, ..., \hat{\beta}_J, ..., \hat{\gamma}_1, ..., \hat{\gamma}_L)$ for the vector of parameters θ, the decision tree can be represented as

$$\hat{y}_{=}\hat{\eta} + \sum_{i=1}^{I} \mu_{A_i}\hat{\alpha}_i + \sum_{j=1}^{J} \mu_{B_j}\hat{\beta}_j + ... + \sum_{l=1}^{L} \mu_{\Gamma_l}\hat{\gamma}_l. \tag{9}$$

Let's consider a situation when as the explaining are used variables, measured in a quantitative scale now. To simplify the presentation, we consider a particular case when the number of input factors is two. We divide the scope of the quantitative variables $x_1,\ x_2$ into fuzzy partitions, which for the first factor we will denote as $A_1,\ A_2, ...,\ A_I$ with the corresponding membership functions $\mu_{1i} \in [0, 1],\ i = \overline{1, I}$. Similarly, for the factor x_2, these will be partitions $B_1,\ B_2, ...,\ B_J$ with membership functions $\mu_{2i} \in [0, 1],\ i = \overline{1, J}$. We will proceed from the fact that at certain sufficiently wide intervals of action of quantitative factors, the behavior of the system's response can be described by a linear dependence. In this case, the complexity of the tree can be reduced by trying to replace the representation of the response in the leaves of the tree, for example, by its linear dependence on input factors. The decision tree from two factors in this case will consist of branches of the form

$$\begin{array}{l} \Pi_{ij}:\ If\ (x_1\ is\ A_i) \wedge (x_2\ is\ B_j)\ \text{then} \\ y'_{ij} = \theta_0 + \theta_{01i} + \theta_{02j} + (\theta_1 + \theta_{11i} + \theta_{12j})x_1 + (\theta_2 x_2 + \theta_{21i} + \theta_{22j})x_2. \end{array} \tag{10}$$

Here, a part of the summands, namely, $\theta_0 + \theta_1 x_1 + \theta_2 x_2$ is included in each branch of the tree and determines the overall linear dependence of the response on input factors throughout the entire area of their definition without taking into account its splitting into partitions. Taking into account (7), the decision tree (10) can be represented as an observation model

$$\begin{array}{l} y_{ijl} = \theta_0 + \sum\limits_{i=1}^{I} \mu_{1i}\theta_{01i} + \sum\limits_{j=1}^{J} \mu_{2j}\theta_{02j} + (\theta_1 + \sum\limits_{i=1}^{I} \mu_{1i}\theta_{11i} + \sum\limits_{j=1}^{J} \mu_{2j}\theta_{12j})x_1 + \\ +(\theta_2 + \sum\limits_{i=1}^{I} \mu_{1i}\theta_{21i} + \sum\limits_{j=1}^{J} \mu_{2j}\theta_{22j})x_2 + e_{ijl}. \end{array} \tag{11}$$

To ensure the identifiability of model (11), we will reduce it by removing a number of regressors from it. For example, you can remove regressors from the model μ_{1I}, μ_{2J}, and also $\mu_{1I}x_1,\ \mu_{1I}x_2,\ \mu_{2J}x_1,\ \mu_{2J}x_2$. Justification of this technique of ensuring the identifiability of the model can be found in [11]- [12].

2 Results and discussions

For the classical linear and polynomial models, many different designs have been proposed (see, for example, , [10], [13]- [17]). As a rule, in order to compare them, they were synthesized on standard domains of input factors, namely, each factor varied in a segment $[-1, +1]$. When considering locally adaptive regression models in the prior construction of optimal designs, we need to make assumptions about the

number, form and location of fuzzy partitions for each factor. At the moment we confine ourselves to the consideration of linear local models. We will assume that in the region of small values of factors, the behavior of the response dependence differs from the behavior of the dependence in the region of large values, for example, in a tilt angle. In the case of consideration of three partitions of such areas with different behavior of the response dependence, there will already be three for each factor. In this paper we will use trapezoidal membership functions. The relative position of the fuzzy partitions will be determined in this case by the coordinates of the intersection points of the neighboring partitions and the width of their intersection zone. Due to the symmetry of the factors domains with respect to zero, we will place fuzzy partitions also symmetrically with respect to a zero value. In the general case, the coordinate of the point of intersection of the partitions will be denoted as $\bar{x}_\mu$. The width of the intersection zone, where the adjacent membership functions have simultaneously non-zero values, will be denoted as 2Δ. The width of the partitions intersection influences the smoothness of the regression dependence transition from one local model to another.

For the classical linear model, the D-optimal experiment design is concentrated at the ends of the segment $[-1, +1]$. These points $+1$ and -1 will be called characteristic. They form the spectrum of the optimal design. For a quadratic model, the point 0 is added to the number of such points. For locally adaptive models with two and three partitions, points with coordinates $\bar{x}_\mu$, $\pm\bar{x}_\mu \pm \Delta$ can be considered as characteristic points. For the case of two partitions, the locally adaptive regression model with one input factor is

$$E(y/x) = f^T(x)\theta = \theta_0 + \theta_1 x + \mu_1(x)\theta_{01} + \mu_1(x)\theta_{11}x \quad ,$$

where $f^T(x) = (1,\ x,\ \mu_1(x),\ \mu_1(x)x)$.

For a case of three partition it takes a form

$$E(y/x) = f^T(x)\theta = \theta_0 + x\theta_1 + \mu_1(x)\theta_{01} + \mu_2(x)\theta_{02} + \mu_1(x)x\theta_{11} + \mu_2(x)x\theta_{12},$$

where $f^T(x) = (1,\ x,\ \mu_1(x),\ \mu_2(x),\ \mu_1(x)x,\ \mu_2(x)x)$.

Tables 1, 2 in the columns entitled as "Spectrum and weights of the $D-$optimal design points" show the weights of the points of the optimal designs spectrum for a linear model with 2 and 3 partitions, and also the values of the information matrix determinant. The designs were constructed using the gradient projection method based on the weights of the points. The accuracy of execution of the necessary and sufficient optimality conditions (5) was provided at the level $\delta = 1 * 10^{-3}$. The designs are symmetrical and the weights of the points symmetric about zero coincide. It is characteristic that the spectra of optimal designs are the characteristic points indicated above.

As can be seen from the analysis of tables 1, 2 at reduction of the width of the intersection of fuzzy partitions efficiency of optimal designs significantly increases. It is quite explainable as the length of the segments on which the membership function for each of the partitions is equal to 1 increases.

Table 1: Characteristics of the D-optimal design for a linear locally adaptive model with two partitions, $\bar{x}_\mu = 0$

Δ	Spectrum and weights of the $D-$optimal design points			$\lvert M(\varepsilon^*)\rvert$
	-1, $+1$	$\bar{x}_\mu - \Delta$, $\bar{x}_\mu + \Delta$	$\bar{x}_\mu$	
0,4	0,2428	0,1941	0,1262	$0,581 * 10^{-3}$
0,3	0,2484	0,2308	0,0416	$0,946 * 10^{-3}$
0,2	0,25	0,25	0	$0,160 * 10^{-2}$
0,1	0,25	0,25	0	$0,256 * 10^{-2}$

Table 2: Characteristics of the D-optimal design for a linear locally adaptive model with three partitions, $\bar{x}_\mu = 0,3$

Δ	Spectrum and weights of the $D-$optimal design points				$\lvert M(\varepsilon^*)\rvert$
	-1, $+1$	$-\bar{x}_\mu - \Delta$, $\bar{x}_\mu + \Delta$	$-\bar{x}_\mu + \Delta$, $\bar{x}_\mu - \Delta$	$-\bar{x}_\mu$, $+\bar{x}_\mu$	
0,2	0,1651	0,1525	0,0893	0,0931	$0,769 * 10^{-7}$
0,15	0,1663	0,1629	0,1505	0,0203	$0,177 * 10^{-6}$
0,1	0,1666	0,1666	0,1666	0	$0,444 * 10^{-6}$
0,05	0,1666	0,1666	0,1666	0	$0,956 * 10^{-6}$

Continuous optimal designs in practice, as a rule, are not applied due to the fact that their full implementation may require a very large number of observations. In practice, discrete optimal designs are applied, which are constructed for a given number of observations. To build them, you can use a combi-gradient algorithm [17] or an algorithm of consecutive addition points in the design [18]- [19]. Due to the fact that, with a relatively small width of the intersection zone, the weights of the points of the spectrum of the continuous D-optimal design are close, it should be expected that the composition of the discrete design points will be mainly formed by the spectrum points of the continuous design. For example, for a model with two partitions and $\Delta = 0,3$ the discrete design with 4 observations included the following points: $\{-1; -0,3; 0,3; 1\}$. The characteristics of this design: $|M(\varepsilon^*)| = 0,937 * 10^{-3}$, $trM^{-1}(\varepsilon^*) = 75,67$, are close to those of the continuous optimal design from table 1.

When considering multifactor models and D-optimal designs for them, the same noted regularities can be observed. Consider for example the case of a two-factor model.

$$E(y/x) = f^T(x)\theta = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \mu_{11}(x_1)\theta_3 + \mu_{21}(x_2)\theta_4 + \mu_{11}(x_1)x_1\theta_5 + \mu_{21}(x_2)x_1\theta_6 + \mu_{11}(x_1)x_2\theta_7 + \mu_{21}(x_2)x_2\theta_8.$$

With a half width of the intersection zone $\Delta = 0,2$ and two partitions, the spectrum of the optimal design will make a full factorial experiment from 16 points with variation levels for each factor $\{-1; -0,2; 0,2; 1\}$. The weights of the points differ depending

on whether they are angular or lie on the edges: the weights of the points of the corner points $(\pm 1; \pm 1)$ are 0,05906, at the points $(\pm 1; \pm 0,2)$ and $(\pm 0,2; \pm 1)$ the weight is 0,07915, at the points $(\pm 0,2; \pm 0,2)$ the weight is 0,03261.

Conclusion

The methodology of a priori experiment design when constructing locally adaptive linear regression models is proposed. The characteristic features of the obtained optimal designs for linear locally adaptive regression models are revealed.

References

[1] Popov A.A., Sautin A.S. (2008) Determination of support vector machines parameters for solving the regression construction problem *Transaction of scientific papers of the Novosibirsk State Technical University* Vol. **2(52)**, pp. 35-40

[2] Popov A.A., Sautin A.S. (2008) Selection of support vector machines parameters for regression using nested grids *The Third International Forum on Strategic Technology, Novosibirsk*, pp. 329-331.

[3] Popov A.A., Boboyev Sh.A. (2015) The construction of a regression relationships using least square in support vector machines *Transaction of scientific papers of the Novosibirsk State Technical University* Vol. **3 (81)**, pp. 69-78.

[4] Takagi T., Sugeno M. (1985) Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. on Systems, Man and Cybernetics* Vol. **15**, No. **1**, pp. 116-132.

[5] Babuska R. (1998) *Fuzzy Modelling for Control.* Kluwer Academic Publishers, London. Boston.

[6] Lilly J.H. (2010) *Fuzzy Control and Identification.* Wiley

[7] Piegat À. (2013) *Fuzzy Modeling and Control.* Beanom publishing house, Moscow.

[8] Popov A.A. (2000) Fuzzy Regression Modeling *Transaction of scientific papers of the Novosibirsk State Technical University* Vol. **2(19)**, pp. 49-57.

[9] Popov A.A., Bykhanov K.V. (2005) Modeling volatility of time series using fuzzy GARCH models *Proceedings - 9th Russian-Korean International Symposium on Science and Technology*, pp. 687-692.

[10] Fedorov V.V. (1971) *Theory of optimal experiments.* Science, Moscow.

[11] Popov A.A. (2009) Construction of decision trees for predicting a quantitative characteristic on a class of logical functions of linguistic variables *Transaction of scientific papers of the Novosibirsk State Technical University* Vol. **3 (36)**, pp. 77-86.

[12] Popov A.A. (1996) Designing discrete and continuous discrete regression type models *Transaction of scientific papers of the Novosibirsk State Technical University* Vol. **1**, pp. 21-30.

[13] Denisov V.I., Popov A.A. (1986) *Optimal experiment design software package.* Finance and Statistics, Moscow.

[14] Krug G.K. Sosulin Yu.A. Fatuev V.A. (1977) *Design of the experiment in identification and extrapolation tasks.* Science, Moscow.

[15] Ermakova S.M. (1983) *Mathematical Theory of optimal experiments.* / Under. ed. S.M. Ermakova (1983) Science, Moscow.

[16] Nalimov V.V., Golikova T.I. (1981) *The logical basis of experiment design.* Metallurgy, Moscow.

[17] Popov A.A. (2013) *Optimal experiment design in problems of structural and parametrical identification of multifactor system models: monograph* NSTU publishing house, Novosibirsk.

[18] Popov A.A. (1995) Sequential schemes for constructing optimal experimental designs *Transaction of scientific papers of the Novosibirsk State Technical University* Vol. **1**, pp. 39-44.

[19] Popov A.A. (2008) Sequential schemes for the synthesis of optimal experimental designs *Proceedings of the Russian higher school academy of sciences* Vol. **1(10)**, pp. 45-55.

On invariant properties of critical Galton-Watson Branching Processes with infinite variance

Azam A. Imomov, Erkin E. Tukhtaev and Navbakhor Nuraliyeva
Karshi State University, Karshi city, Uzbekistan
e-mail: imomov_azam@mail.ru, tukhtaev_erkin@mail.ru, n.navbahor@mail.ru

Abstract

We observe the Galton-Watson Branching Processes. Limit properties of transition functions and their convergence to invariant measures are investigated.

Keywords: Branching process, Immigration, Transition probabilities, Slow variation, Invariant measures.

Introduction

Let $\{X_n, n \in \mathbb{N}_0\}$ be the Galton-Watson Branching Process allowing Immigration (GWPI), where $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$ and $\mathbb{N} = \{1, 2, \ldots\}$. This is a homogeneous Markov chain with state space $\mathcal{S} \subset \mathbb{N}_0$ and whose transition probabilities are

$$p_{ij} = \text{coefficient of } \ s^j \ \text{ in } \ h(s)\big(f(s)\big)^i, \quad s \in [0, 1),$$

where $h(s) = \sum_{j\in\mathcal{S}} h_j s^j$ and $f(s) = \sum_{j\in\mathcal{S}} p_j s^j$ are probability generating functions (PGF's). The variable X_n is interpreted as the population size in GWPI at the moment n. An evolution of the process will occurs by following scheme. An initial state is empty that is $X_0 = 0$ and the process starts owing to immigrants. Each individual at time n produces j progeny with probability p_j independently of each other so that $p_0 > 0$. Simultaneously in the population i immigrants arrive with probability h_i in each moment $n \in \mathbb{N}$. These individuals undergo further transformation obeying the reproduction law $\{p_j\}$ and n-step transition probabilities $p_{ij}^{(n)} := \mathsf{P}\{X_{n+k} = j | X_k = i\}$ for any $k \in \mathbb{N}$ are given by

$$\mathcal{P}_n^{(i)}(s) := \sum_{j\in\mathcal{S}} p_{ij}^{(n)} s^j = \big(f_n(s)\big)^i \prod_{k=0}^{n-1} h\big(f_k(s)\big) \quad \textit{for any} \ \ i \in \mathcal{S}, \tag{1}$$

where $f_n(s)$ is n-fold iteration of PGF $f(s)$; see for example [7]. Thus the transition probabilities $\left\{p_{ij}^{(n)}\right\}$ are completely defined by the probabilities $\{p_j\}$ and $\{h_j\}$.

Classification of states of the chain $\{X_n\}$ is one of fundamental problems in theory of GWPI. Direct differentiation of (1) gives

$$\mathsf{E}\big[X_n | X_0 = i\big] = \begin{cases} an + i & , \quad \textit{when} \ \ m = 1, \\ \left(\dfrac{a}{m-1} + i\right) m^n - \dfrac{a}{m-1}\,, & \quad \textit{when} \ \ m \neq 1, \end{cases}$$

where $m = f'(1-)$ is mean per-capita offspring number and $a = h'(1-)$. The received formula for $\mathsf{E}\big[X_n | X_0 = i\big]$ shows that classification of states of GWPI depends on a value of the parameter m. Process $\{X_n\}$ is classified as sub-critical, critical and supercritical if $m < 1$, $m = 1$ and $m > 1$ accordingly.

The above described population process was considered first by Heathcote [3] in 1965. Further long-term properties of $\mathcal{S}$ and a problem of existence and uniqueness of invariant measures of GWPI were investigated by Seneta [13], Pakes [9], [10] and by many other authors. Therein some moment conditions for PGF $f(s)$ and $h(s)$ was required to be satisfied. In aforementioned works of Seneta the ergodic properties of $\{X_n\}$ were investigated. He has proved that when $m \leq 1$ the process $\{X_n\}$ has an invariant measure $\{\mu_k,\ k \in \mathcal{S}\}$ which is unique up to multiplicative constant. Pakes [10] have shown that in supercritical case $\mathcal{S}$ is transient. In the critical case $\mathcal{S}$ can be transient, null-recurrent or ergodic. In this case, if in addition to assume that $2b := f''(1-) < \infty$, properties of $\mathcal{S}$ depend on value of parameter $\lambda = a/b$: if $\lambda > 1$ or $\lambda < 1$, then $\mathcal{S}$ is transient or null-recurrent accordingly. In the case when $\lambda = 1$, Pakes [9] studied necessary and sufficient conditions for a null-recurrence property. Limiting distribution law for critical process $\{X_n\}$ was found first by Seneta [12]. He has proved that the normalized process $X_n/(bn)$ has limiting Gamma distribution with density function $\Gamma^{-1}(\lambda)x^{\lambda-1}e^{-x}$ provided that $0 < \lambda < \infty$, where $x > 0$ and $\Gamma(*)$ is Euler's Gamma function. This result has been established also by Pakes [9] without reference to Seneta. Afterwards Pakes [7], [8], has obtained principally new results for all cases $m < \infty$ and $b = \infty$.

Throughout the paper we keep on the critical case only and $b = \infty$. Our reasoning will bound up with elements of slow variation theory in sense of Karamata; see [11]. Remind that real-valued, positive and measurable function $L(x)$ is said to be slowly varying (SV) at infinity if $L(\lambda x)/L(x) \to 1$ as $x \to \infty$ for each $\lambda > 0$. We refer the reader to [1], [2] and [11] for more information.

In second section we study invariant measures of the simple Galton-Watson (GW) Process. In third section the invariant properties of GWPI will be investigated.

1 Invariant measures of GW Process

Let $\{Z_n, n \in \mathbb{N}_0\}$ be the simple GW Branching Process without immigration given by offspring PGF $f(s)$. Discussing this case we will assume that the offspring PGF $f(s)$ has the following representation:

$$f(s) = s + (1-s)^{1+\nu}\mathcal{L}\left(\frac{1}{1-s}\right), \qquad [f_\nu]$$

where $0 < \nu \leq 1$ and $\mathcal{L}(x)$ is SV at infinity. By the criticality of the process the condition $[f_\nu]$ implies that $b = \infty$. This includes the case $b < \infty$ when $\nu = 1$ and $\mathcal{L}(t) \to b$ as $t \to \infty$.

Consider PGF $f_n(s) := \mathsf{E}\left[s^{Z_n} | Z_0 = 1\right]$ and write $R_n(s) := 1 - f_n(s)$. Evidently $Q_n := R_n(0)$ is the survival probability of the process. By arguments of Slack [14]

one can be shown that if the condition $[f_\nu]$ holds then

$$Q_n^\nu \cdot \mathcal{L}\left(\frac{1}{Q_n}\right) \sim \frac{1}{\nu n} \quad as \quad n \to \infty. \tag{2}$$

Slack [14] also has shown that

$$\mathcal{U}_n(s) := \frac{f_n(s) - f_n(0)}{f_n(0) - f_{n-1}(0)} \longrightarrow U(s) \tag{3}$$

for $s \in [0, 1)$, where the limit function $U(s)$ satisfies the Abel equation

$$U\left(f(s)\right) = U(s) + 1, \tag{4}$$

so that $U(s)$ is PGF of invariant measure for the GW process $\{Z_n\}$. Combining $[f_\nu]$, (2) and (3) and considering properties of the process $\{Z_n\}$ we have

$$\mathcal{U}_n(s) \sim U_n(s) := \left[1 - \frac{R_n(s)}{Q_n}\right] \nu n \quad as \quad n \to \infty.$$

So we proved the following lemma.

Lemma 1. *If the condition* $[f_\nu]$ *holds then*

$$R_n(s) = \frac{\mathcal{N}(n)}{(\nu n)^{1/\nu}} \cdot \left[1 - \frac{U_n(s)}{\nu n}\right], \tag{5}$$

where the function $\mathcal{N}(x)$ *is SV at infinity and*

$$\mathcal{N}(n) \cdot \mathcal{L}^{1/\nu}\left(\frac{(\nu n)^{1/\nu}}{\mathcal{N}(n)}\right) \longrightarrow 1 \quad as \quad n \to \infty, \tag{6}$$

and the function $U_n(s)$ *enjoys following properties:*

- $U_n(s) \longrightarrow U(s)$ *as* $n \to \infty$ *so that the equation* (4) *holds;*
- $\lim_{s\uparrow 1} U_n(s) = \nu n$ *for each fixed* $n \in \mathbb{N}$*;*
- $U_n(0) = 0$ *for each fixed* $n \in \mathbb{N}$*.*

Evidently that this lemma is generalization of (2) and herein it established by more simple proof rather than as shown in [4].

Further writing $\Lambda(y) = y^\nu \mathcal{L}\left(1/y\right)$ we consider the function

$$\mathcal{M}_n(s) := 1 - \frac{\Lambda\big(R_n(s)\big)}{\Lambda\left(Q_n\right)}. \tag{7}$$

It follows from (5) and from the properties of SV-function that

$$\begin{aligned}\mathcal{M}_n(s) &= 1 - \left(\frac{R_n(s)}{Q_n}\right)^\nu \frac{\mathcal{L}\big(1/R_n(s)\big)}{\mathcal{L}\big(1/Q_n\big)} \\ &\sim 1 - \left(1 - \frac{U_n(s)}{\nu n}\right)^\nu \sim \frac{U_n(s)}{n}\big(1 + \rho_n(s)\big) \quad as \quad n \to \infty,\end{aligned}$$

where $\rho_n(s) = \mathcal{O}\big(1/n\big)$ uniformly for all $s \in [0, 1)$.

Thus we obtain the following assertion.

Lemma 2. *If the condition* $[f_\nu]$ *holds then*

$$n \cdot \mathcal{M}_n(s) \longrightarrow U(s) \quad as \quad n \to \infty, \tag{8}$$

where $U(s)$ *is PGF of invariant measure of GW Process.*

In the following theorem we find an explicit form of PGF $U(s)$. Write

$$\mathcal{V}(s) := \frac{1}{\nu\Lambda\,(1-s)}.$$

Theorem 1. *If the condition* $[f_\nu]$ *holds then*

$$U(s) = \mathcal{V}(s) - \mathcal{V}(0). \tag{9}$$

Proof. In pursuance of reasoning from [2, p. 401] we obtain the following relation:

$$\mathcal{V}\big(f_{n+1}(s)\big) - \mathcal{V}\big(f_n(s)\big) \longrightarrow 1 \quad as \quad n \to \infty.$$

Thence summing by n we find

$$\mathcal{V}\big(f_n(s)\big) - \mathcal{V}(s) = n \cdot \big(1 + o(1)\big) \quad as \quad n \to \infty.$$

Keeping our designation we easily will transform last equality to a form of

$$\Lambda\big(R_n(s)\big) = \frac{\Lambda\,(1-s)}{\Lambda\,(1-s)\,\nu n + 1}\big(1 + o(1)\big) \quad as \quad n \to \infty. \tag{10}$$

Combining (7), (8) and (10) we reach (9). □

2 Invariant measures of GWPI

Consider GWPI. Pakes [8] has proved the following theorem.

Theorem P1 ([8]). *If* $m = 1$ *then*

$$p_{00}^{(n)} \sim K \exp\left\{\int\limits_1^{e^n} \frac{\ln h\,(1-\varphi(y))}{y}\,dy\right\} \quad as \quad n \to \infty,$$

where $\varphi(y)$ *is decreasing SV-function. If*

$$\sum_{m=0}^{\infty}\Big[(1 - h(f_m(0))\,(1 - f'(f_m(0))\Big] < \infty,$$

then

$$p_{00}^{(n)} \sim K_1 \exp\left\{\int\limits_0^{f_n(0)} \frac{\ln h(y)}{f(y) - y}\,dy\right\} \quad as \quad n \to \infty.$$

Herein K and K_1 are some constants.

Since this point we everywhere will consider the case that immigration PGF $h(s)$ has the following form:

$$1-h(s)=(1-s)^{\delta}\ell\left(\frac{1}{1-s}\right), \qquad [h_\delta]$$

where $0<\delta\le 1$ and $\ell(x)$ is SV at infinity.

Our results appear provided that conditions $[f_\nu]$ and $[h_\delta]$ hold and $\delta>\nu$. As it has been shown in [8] that in this case $\mathcal{S}$ is ergodic. Namely we improve statements of Theorem P1. Herewith we put forward an additional requirement concerning $\mathcal{L}(x)$ and $\ell(x)$. So since $\mathcal{L}(x)$ is SV we can write

$$\frac{\mathcal{L}\,(\lambda x)}{\mathcal{L}(x)}=1+\alpha(x) \qquad [\mathcal{L}_\alpha]$$

for each $\lambda>0$, where $\alpha(x)\to 0$ as $x\to\infty$. Henceforth we suppose that some positive function $g(x)$ is given so that $g(x)\to 0$ and $\alpha(x)=o\big(g(x)\big)$ as $x\to\infty$. In this case $\mathcal{L}(x)$ is called SV with remainder $\alpha(x)$; see [2, p. 185, condition SR3]. Wherever we exploit the condition $[\mathcal{L}_\alpha]$ we will suppose that

$$\alpha(x)=o\left(\frac{\mathcal{L}\,(x)}{x^{\nu}}\right)\quad as\quad x\to\infty. \tag{11}$$

And also by perforce we suppose the condition

$$\frac{\ell\,(\lambda x)}{\ell(x)}=1+\beta(x) \qquad [\ell_\beta]$$

for each $\lambda>0$, where

$$\beta(x)=o\left(\frac{\ell\,(x)}{x^{\delta}}\right)\quad as\quad x\to\infty.$$

Since $f_n(s)\uparrow 1$ for all $s\in[0,1)$ in virtue of (1) it sufficiently to observe the case $i=0$ as $n\to\infty$. Write

$$\mathcal{P}_n(s)=\mathcal{P}_n^{(0)}(s).$$

The following theorem is generalization of the Theorem P1.

Theorem 2. *Let conditions $[f_\nu]$, $[h_\delta]$ hold. If $\delta>\nu$ then*

$$\mathcal{P}_n(s)\sim K(s)\exp\left\{-\int\limits_s^{f_n(s)}\frac{1-h(y)}{f(y)-y}\Big[1+\delta(1-y)\Big]dy\right\}$$

as $n\to\infty$, where $K(s)$ is a bounded function for $s\in[0,1)$ and $\delta(x)\to 0$ as $x\downarrow 0$. If in addition, the conditions $[\mathcal{L}_\alpha]$ and (11) are satisfied then

$$\mathcal{P}_n(s)\sim K(s)\exp\left\{-\int\limits_s^{f_n(s)}\frac{1-h(y)}{f(y)-y}\Big[1+o\big(\Lambda\,(1-y)\big)\Big]dy\right\}\quad as\quad n\to\infty.$$

Corollary 1. *Let conditions* $[f_\nu]$, $[h_\delta]$ *hold. If* $\delta > \nu$ *then*

$$p_{00}^{(n)} \sim A \exp\left\{-\frac{\mathcal{N}^\nu(n)}{\delta-\nu}\cdot \ell\left(\frac{(\nu n)^{1/\nu}}{\mathcal{N}(n)}\right)\right\} \quad as \quad n\to\infty,$$

where A *is a positive constant and* $\mathcal{N}(x)$ *is SV at infinity defined in* (6).

Further we need the following result which is an improved analog of the Basic Lemma of the theory of critical GW processes.

Lemma 3 ([5]). *Let conditions* $[f_\nu]$, $[\mathcal{L}_\alpha]$ *and* (11) *hold. Then*

$$\frac{1}{\Lambda\big(R_n(s)\big)} - \frac{1}{\Lambda\,(1-s)} = \nu n + \frac{1+\nu}{2}\cdot \ln\left(1+\nu n\Lambda(1-s)\right) + \rho_n(s),$$

where $\rho_n(s) = o\big(\ln n\big) + \sigma_n(s)$ *and,* $\sigma_n(s)$ *is bounded uniformly for* $s\in[0,1)$ *and converges to a limit* $\sigma(s)$ *as* $n\to\infty$ *which is a bounded function of* $s\in[0,1)$.

We make sure that at the conditions of second part of Theorem 2 PGF $\mathcal{P}_n(s)$ converges to a limit $\pi(s)$ which we denote by the power series representation

$$\pi(s) = \sum_{j\in\mathcal{S}} \pi_j s^j.$$

Using the Lemma 3 we will establish a speed rate of this convergence in the following theorem.

Theorem 3. *Let conditions* $[f_\nu]$, $[h_\delta]$ *hold and* $\delta > \nu$. *Then* $\mathcal{P}_n(s)$ *converges to* $\pi(s)$ *which generates the invariant measures* $\{\pi_j\}$ *for GWPI. The convergence is uniform over compact subsets of the open unit disc. If in addition, the conditions* $[\mathcal{L}_\alpha]$, (11) *and* $[\ell_\beta]$ *are fulfilled then*

$$\mathcal{P}_n(s) = \pi(s)\left(1+\Delta_n(s)\mathcal{N}_\delta\left(\frac{1}{R_n(s)}\right)\right),$$

where $\mathcal{N}_\delta(x) = \mathcal{N}^\delta(x)\ell(x)$, *the function* $\mathcal{N}(x)$ *is defined in* (6) *and*

$$\Delta_n(s) = \frac{1}{\delta-\nu}\,\frac{1}{\big(\nu_n(s)\big)^{\delta/\nu-1}} - \frac{1+\nu}{2\nu}\,\frac{\ln\big[\nu_n(s)\big]}{\big(\nu_n(s)\big)^{\delta/\nu}}\big(1+o(1)\big)$$

as $n\to\infty$ *and* $\nu_n(s) = \nu n + \Lambda^{-1}\,(1-s)$.

The following result is direct consequence of Theorem 3.

Corollary 2. *If conditions of Theorem 3 hold then*

$$p_{00}^{(n)} = \pi_0\cdot\Big(1+\Delta_n\mathcal{N}_\delta\,(n)\Big),$$

where $\mathcal{N}_\delta(n)$ *is SV at infinity and*

$$\Delta_n = \frac{1}{\delta-\nu}\,\frac{1}{(\nu n)^{\delta/\nu-1}} - \frac{1+\nu}{2\nu}\,\frac{\ln n}{(\nu n)^{\delta/\nu}}\big(1+o(1)\big) \quad as \quad n\to\infty.$$

Remark 1. *The analogous result as in Theorem 3 has been proved in [6] provided that* $\delta = 1$ *and* $f'''(1-) < \infty$.

References

[1] Asmussen S., Hering H. (1983). *Branching processes*. Birkhäuser, Boston.

[2] Bingham N. H., Goldie C. M., Teugels J. L. (1987). *Regular Variation*. Univ. Press, Cambridge.

[3] Heatcote C. R. (1965). A branching process allowing immigration. *Jour. Royal Stat. Soc.* Vol. **B-27**, pp. 138–143.

[4] Imomov A. A. (2019). On a limit structure of the Galton-Watson branching processes with regularly varying generating functions. *Prob. and Math. Stat.*, Vol. **39(1)**, pp. 61–73.

[5] Imomov A. A., Tukhtaev E. E. (2019). On Application of Slowly Varying Functions with Remainder in the Theory of Galton-Watson Branching Process. *J Siber. Fed. Univ.: Math. Phys.* Vol. **12(1)**, pp. 51–57.

[6] Imomov A. A. (2015). On long-time behaviors of states of Galton-Watson Branching Processes allowing Immigration. *J Siber. Fed. Univ.: Math. Phys.* Vol. **8(4)**, pp. 394–405.

[7] Pakes A. G. (1979). Limit theorems for the simple branching process allowing immigration, I. The case of finite offspring mean. *Adv. Appl. Prob.* Vol. **11**, pp. 31–62.

[8] Pakes A. G. (1975). Some results for non-supercritical Galton-Watson process with immigration. *Math. Biosci.* Vol. **24**, pp. 71–92.

[9] Pakes A. G. (1971). On the critical Galton-Watson process with immigration. *Jour. Austral. Math. Soc.* Vol. **12**, pp. 476–482.

[10] Pakes A. G. (1971). Branching processes with immigration. *Jour. Appl. Prob.* Vol. **8(1)**, pp. 32–42.

[11] Seneta E. (1972). *Regularly Varying Functions*. Springer, Berlin.

[12] Seneta E. (1970). An explicit-limit theorem for the critical Galton-Watson process with immigration. *Jour. Royal Stat. Soc.* Vol. **B-32(1)**, pp. 149–152.

[13] Seneta E. (1969). Functional equations and the Galton-Watson process. *Adv. Appl. Prob.* Vol. **1**, pp. 1–42.

[14] Slack R. S. (1972). Further notes on branching processes with mean 1. *Wahrscheinlichkeitstheor. und Verv. Geb.* Vol. **25**, pp. 31–38.

On some practical approaches of data science applied in forecasting and personalization

Milovan Krnjajić[1] and Roman Maslovskis[2]
[1] *Maxim Integrated Products, San Jose, CA, USA*
[2] *eBay Inc, San Jose, CA, USA*
e-mail: krnjajic@yahoo.com, rmaslovskis@ebay.com

Abstract

Data scientists use statistical models and methods along with algorithmic (machine-learning) approaches to solve problems of classification, forecasting, pattern recognition, inference and interpretation of results. Practical difficulties include dealing with enormous datasets with complex structures requiring substantial computational support. Moreover, when applied in real-world business, the challenges for data science multiply as practice induces an additional level of complexity requiring solutions optimized with respect to cost, time, and other specific regulatory, financial or environmental constraints. Here it is important to complement scientific methods with common sense approach, practical heuristics and judicious decision making during all important phases of developing data driven solutions. We focus on two complementary yet representative data science topics: (1) a problem of periodic automated sales forecasting for a large number of industrial products where we survey forecasting models and methods along with practical procedures for evaluation of their predictive performance and weigh empirical evidence of their relative merits in model selection, and (2) a problem of cluster analysis as a basis for personalization and segmentation of customers and products in a typical e-commerce application.

Disclaimer: This article expresses personal views and opinions of the authors, which may not coincide with policies or positions of their employers. Examples were chosen for the purpose of illustrating main ideas and do not refer necessarily or directly to their work on any specific project.

Keywords: forecasting intermittent demand, personalization, segmentation, clustering, angular distance measure.

Introduction

Companies which develop, manufacture or sell thousands of products need regular and automated forecasts to support various functions such as supply chain management, sales operations and financial planning. Forecasts of demand for products/parts are typically done on monthly or weekly basis for multiple horizons from one month to a year. Predictions are based on analysis of individual time-series of various lengths depending on product age. Given the number of time-series in tens of thousands or more, automating forecasts becomes a necessity. We do not consider commercial forecasting packages but work with those developed as a result of mainly academic research and implemented on freely available software platforms such as CRAN's R programming environment, see [14], for example.

Traditional approach to forecasting mostly uses statistical time-series models, however, along with re-emergence of neural networks as useful non-parametric tools for pattern recognition, more machine learning methods have been used for classification and prediction. As is usual with heterogeneous methodologies, each of these approaches may have advantages over the other depending on concrete data and application area. A number of empirical studies, such as one presented in [11] have been conducted and practical evidence gathered comparing relative merits of two approaches in terms of predictive accuracy and requirements for computational resources.

When evaluating predictive performance and computing requirements, in order to decide which forecasting methodology to use, it is important to consider complexity of methods and interpretability of the results, as advocated in [6]. Availability of substantial computational power allows for checking of a multitude of forecasting models and methods (off-the-shelf ones or those developed for specific applications) thus making explicit the uncertainty inherently present in the model selection process. Model choice is usually based on a series of evaluations of individual predictive performance, with a need to employ some form of regularization to counter over-fitting to the selection procedure and so avoid selection bias, as shown in [3]. Another important ingredient in performance estimation is choice of the right accuracy metric, which is especially critical in case of intermittent demand, see [7] and [4].

Complementary approach to searching for a "best" model is to combine forecasts from a number of models as first proposed in [2]. Combined forecasts have some appealing properties in terms of the upper bound of the predictive error. Importantly, predictive performance for a large number of individual datasets can also be improved by aggregating time series in groups and hierarchies and reconciling forecasts at different levels as presented in [16], and specifically for intermittent demand in [12].

The first section provides a brief survey of select models, methods and procedures which we find practical and effective in developing forecasting solutions on industrial scale for the real-world problems. Our preferences for certain approaches are supported by published empirical evidence from diverse forecasting applications.

The second part of the paper focuses on segmentation and clustering as a basis for online store personalization in modern e-commerce. The goal of personalization is to offer merchandise desired by customers without need for their explicit querying for it. Techniques for personalization leverage all available information about customers and product characteristics along with history of purchases and detailed interaction with the e-commerce web-site. The processes and methods providing personalization can be viewed as an application of data mining techniques [10].

At the core of the methods enabling personalized online views are clustering methods which segment the spaces of customers and products into associated groups. Depending on the shape of clusters, proximity measures can result in very different groupings which in turn may poorly reflect the true classification. Thus, one of the challenges of applying clustering algorithms is the selection of an appropriate proximity (distance) measure.

We illustrate clustering examples in the online space of products and buyers where

Euclidean distance works well in one class of cases but yields unsatisfactory associations in the other, whereas the cosine measure has opposite performance. As a solution we introduce a novel proximity measure, angular distance measure (ADM), which performs well in both cases but also in a case where neither Euclidean nor cosine measure is satisfactory.

1 Forecasting for a large number of time series

A company engaged in research and development (R&D), manufacturing or sales of thousands of different products or parts (for example, auto parts or electronic circuits) needs forward looking capabilities in order to make adequate (1) planning of production capacity, (2) management of supply chain, (3) financial planning, (4) deployment of sales force and (5) directing efforts of marketing organization. All of these functions and processes use forecasting projections of the future demand for parts, which means that applying reliable forecasting procedures is a condition of critical importance for business success. Our goal is to identify, select, develop and apply effective predictive approaches and methodologies based on objective scientific criteria and proven industrial practices.

1.1 Difficulties with intermittent demand

Analyzing sales of parts/products implies working with time-series that may experience multiple periods of zero demand. Intermittent demand (ID) is a common place in many industries from aviation and automotive to defense and semiconductor and also occurs with products approaching the end of their life cycle. In case of analyzing tens of thousands of parts a large fraction may be of intermittent demand, what has direct consequences for inventory control and maintenance of stock levels. Clearly, reliably and accurately forecasting irregular demand has critical importance for optimizing inventory levels, with direct financial impact on the business.

Difficulty with ID time-series is that usual forecasting methods, such as simple exponential smoothing (SES), may not apply. Variants of SES compute a forecast for the next period as a weighted average, $F_{t+1} = \alpha Y_t + (1-\alpha)F_t$, of the the current forecast, F_t, and demand, Y_t, and exhibit an upward bias for periods after zero-demand. More robust forecasts are obtained using Croston's method and its derivatives (see [6], and references therein) although these are not based on statistical models. Briefly, Croston's method is based on separate estimates of demand size and time-intervals. Adjusted versions estimate the probability of non-zero demand (instead of the interval size), use additional parameters, and lower the bias of estimates.

Model based forecasts for ID include flavors of auto-regressive moving average (ARMA) models such as discrete and integer-valued ARMA which can describe stationary stochastic processes taking non-negative integer values. Alternative methods use bootstrapping and temporal aggregation of time series which in case of sufficiently large data may remove zeroes and allow for predictions to be done using standard methods such as SES.

Petropoulos et al. in [12] describe an approach to forecasting ID using multiple aggregated views of the data (e.g., daily observations turned into weekly, bi-weekly or monthly). Temporal aggregation lessens intermittence, however, it results in a loss of information and it's also difficult to determine an "optimal" aggregation level. A solution is to combine forecasts made on multiple temporal aggregation levels resulting in improved predictions as shown empirically in the same paper. A complementary approach to temporal aggregation is described in [13] and is based on transforming the time series, by switching the roles of time and demand, to represent inter-demand intervals per cumulative demand. Transformed data have smaller variance and predictions show improved accuracy over analysis with original data.

1.2 Automated forecasting procedures and accuracy metrics

Making forecast for tens of thousands of time series would be next to impossible without some kind of automation of the process. Commercial packages such as SAS and Autobox have capability for automated forecasting, whereas freely available CRAN-R platform offers libraries of functions which can be combined into a flexible forecasting system for specific needs. We worked with `forecast` package, developed by R. Hyndman et al., see [8].

The package implements automatic forecasts using a number of methods and models including exponential smoothing, ARIMA models, Theta method, and splines. We worked with exponential smoothing methods (ESM) and ARIMA models. It should be noted that ESM-s are algorithms for producing point forecasts only, but reformulating them in terms of underlying state space stochastic models provides a framework for computing prediction intervals along with the same point forecasts. Akaike Information Criterion (AIC) is used for model selection. Non-linear ESM-s are intended for non-stationary series, whereas ARIMA models work better for stationary data. Automatic selection of ARIMA models is done through `auto.arima()` function based on a combination of unit root tests, MLE and minimization of the AIC.

In general, automatic models selection is not without risk, since spurious relationships are bound to happen in a multi-testing setting. Moreover, overfitting the training data, especially in case of shorter time-series, is another problem that may adversely affect predictive performance of the chosen model. Performance of forecasting methods or models needs to be evaluated in order to measure how predictions compare to the actually observed values. Model predictions performance computed on a set of observed points not used in model fitting is then used as the criterion for selecting a preferred model. It is therefore important to choose appropriate accuracy metrics as described in [7].

Typical forecast accuracy metrics compute mean absolute error (MAE), or its variants using median or geometric mean. These errors are on the same scale as the data so accuracy measurements are scale-dependent. The problem, especially for intermittent demand, is that output of these functions may be infinite or undefined.

Scale independence is provided by mean absolute percentage error (MAPE) and this is one of the widely used accuracy metrics in demand planning, for example.

However, MAPE, defined as the mean of absolute values $100(Y_i - F_i)/Y_i$, may also produce undefined or infinite results on the account of division by zero. Moreover, distribution of errors can be very skewed in case when actual values are close to zero. Another problem with MAPE is that it favors smaller rather than larger forecasts which can be avoided using the symmetric version of MAPE.

MASE is the mean of the absolute values of scaled errors q_t defined as the ratio between $\epsilon_t = Y_t - F_t$ and the prediction of a naive method such as average one-step forecast, $\sum_{j=2}^{n} |Y_j - F_j|/(n-1)$. Mean absolute scaled error (MASE) is scale independent and is always defined except when all observations are the same. In addition to MASE, working with the error measure based on $\log(F_i/Y_i)$ is recommended as it has desirable statistical properties, see [15].

1.3 ML vs time-series models

Modern methods of data analysis are based on automated algorithms for solving large classes of problems without the need for explicit encoding of specific solutions. This approach goes by names of artificial intelligence (AI), machine learning (ML) or deep learning (DL) and has gained traction and prominence in recent years in diverse areas of speech recognition, language translation, computer vision, robotics, games, and many other applications in pattern recognition, classification and prediction.

Typical ML methods include various versions of neural networks, tree-based methods (such as random forests, gradient boosting machines), kernel methods, and Bayesian networks. The successes of these methods is largely based on using of algorithms capable of gradual learning of the underlying structures in the data by trial and error and iterative improvement of the performance.

We've found flexibility of ML methods to be valuable when the main goal is to classify data points without a need to interpret the results. In particular, tree-based methods such as random forests (RF) proved to be most useful for classification tasks given their (almost) plug-and-play property as it doesn't require extensive optimization of hyper-parameters. For example, when forecasting intermittent demand for a specific time-horizon, time-series may be classified in two groups, one of which is with predicted zero demand, and then these results combined with the forecast from time-series methods.

Large research body of literature exists on advances of new ML/AI methods [5], with growing number of applications of ML approaches to forecasting problems, especially using long-short-term-memory (LSTM), NN-s and tree based algorithms. Yet, limited objective assessment of performance of ML methods in forecasting is available to help guide practitioners in deciding when applying ML techniques may be advantageous over traditional, statistical methods and time-series models.

Typically, academic papers on ML forecasting report forecasts performance but do not provide comparison with simple statistical methods or naive benchmarks, and so imply that ML methods make superior predictions although not providing any supporting empirical evidence. This resembles the situation in statistical literature in the 1970s and 1980s when forecasting methods were considered to have superior

predictive accuracy simply because of model complexity and mathematical appeal.

In a recent comparative analysis Makridakis et al. [11] evaluate forecasting performance of ML methods and statistical models on a rich set (of 1000^{+}) of time series. The study shows that simple time-series models such as linear methods, exponential smoothing, and ARIMA outperform complex ML approaches, including Multi-Layer Perceptron (MLP), Bayesian NNs (BNN), Kernel Regression NNs, Radial Basis Functions (RBF), Classification and Regression trees (CART), Gaussian Processes (GP), and Support Vector Regression (SVR). ML methods also are more computationally intensive and in general require more data along with feature engineering.

1.4 Temporal aggregating and hierarchical forecasting

Typical business requires decision making at different levels of activity and also needs forecasts for multiple time horizons. Strategic considerations use aggregated long-run forecasts whereas operational decisions that are highly dynamical and time constrained require detailed individual and short-term forecasts.

Forecasts are done separately at each temporal level using different approaches and are based on multiple sources of information so it's expected that the results may not agree. Yet, for consistent decision making it is important to reconcile the aggregated predictions at multiple temporal and aggregation levels.

Athanasopoulos et al. in [1] introduce a concept of temporal hierarchies in forecasting of time-series and the method is implemented in the `thief` R package. Temporal hierarchy is represented by the connections of grouping structures across the aggregation levels. Main advantage of using temporal hierarchies is in merging data from different sources and optimally combining predictions from different aggregation levels into consistent and reconciled forecasts. This combination allows for better planning and decision making and also improving forecasting accuracy. Reconciliation across aggregation levels ensures that higher level forecasts are equal to the sums of subaggregate predictions. Reconciliation is optimal in the sense that predictions at individually forecasted levels are adjusted by the least amount to produce consistent values across the hierarchy.

A complementary approach to working with temporal hierarchies is to forecast on hierarchies of time series disaggregated by various attributes. For example, total count of manufactured individual parts can be disaggregated into groups by product function or type. These categories can be nested within larger groups so that the collection of corresponding time-series would follow a hierarchical aggregation pattern. Additional to product hierarchy, time series can be grouped according to geography of sales to different countries and regions. Such aggregation produces more general structures which are based on both nested and crossed disaggregation factors. Consistent forecasting requires the forecasts to add up in the same way as the data. For example, forecasts for individual parts should add up to forecasts across grouping hierarchies. Traditional ways of reconciling forecasts at different levels are not satisfactory since bottom-up approach, for example, starts with noisy bottom-level data with errors compounding, whereas the top-down disaggregation introduces bias. A

recent approach, described in [16] and implemented in `hts` R package, avoids above disadvantages and produces consistent forecasts by minimizing the mean squared errors across the entire collection of time series under the assumption of unbiasedness.

1.5 Search for a "best" model and combining forecasts

A trend toward use of more complex forecasting methods has been observed in previous few decades with econometricians setting the trend (see for example [6] and references therein). Yet it has been shown in a number of empirical studies (not only in case of ML vs traditional time-series methods) that simple models were at least as accurate and often more accurate than complex models. Here, simplicity of the model is not understood in terms of the number of variables in regression model, or amount of model development effort, for example. Instead, simple forecasting model/method is taken to be a forecasting process that is *understandable* to forecast users.

Reluctance to trust and use simple methods comes from bias towards complexity, which is deemed more persuasive, as explicitly shown in experiments using abstracts from academic papers – abstracts containing an algebraic equation were judged to be of higher quality although the mathematical expression was unrelated to the contents.

Fulfilling the goals of the exercise to make sensible forecasts for many thousands of time-series has a form of a giant optimization which includes repeated searching in the space of possible models for those that are "best". An example of one level of this search is obtaining optimal ARIMA model as implemented within `auto.arima()` function of the `forecast` package and using a procedure based on computing unit roots, MLE and AIC to estimate optimal model parameters. Another example is searching the space of such "optimal" models in attempt to select a subset of those with the "best" predictive performance evaluated on the test set not used in model fitting (and calculating MLE and AIC). Some properties of principled model selection based on comparison of predictive performance are described in [9].

In our approach to forecasting for a large number of time-series we prefer using a number of different yet straightforward procedures and methods. We (1) select automatically time series models with optimal parameters (`forecast` R-package and `auto.arima()` function), (2) use hierarchical multi-horizon methods, and (3) combine predictive outputs of a select subset of "best" methods. However, it is difficult, if not impossible, to find a single model outperforming all others on multiple datasets in a specific application. It has been suggested decades ago [2] that combined forecasts of several models or methods outperform, on average, each individual approach in terms of predictive accuracy. Similarly, Bayesian model averaging allows for weighted combination of models' outputs. Combining forecasts has advantage of reducing variability of prediction errors, and in the case of two models, upper bound for the error variance of the combined forecast is the smaller of the individual variances.

In addition, we partition the set of time series in natural groups (based on similarity of product, for example) and apply classes of multiple forecasting models and methods to each group. Predictive accuracy of forecasts is always evaluated on the observations not used in model fitting and then the outputs of a subset of best per-

forming methods are combined as simple (or weighted) averages into the final forecast.

Forecasts obtained from this procedure can also be viewed as exhibiting a selection bias or, in some sense, overfitting this very process as described in [3]. The authors suggest a form of a cross-validation protocol which requires model selection to be an integral part of the model fitting process with selection step performed anew whenever a model is trained with a new set of data. Details are too many to explain, but we are confident that we have largely followed the main trust of this approach.

2 Clustering of consumer preferences in e-commerce

A typical modern e-commerce site deploys systems which attempt to treat each customer in a "personalized" way so as to offer the content that is likely to be of interest to the user. Specifically, in developing a personalized system for an online store the goal is to identify and offer those groups of products that are in line with customer preferences and so has a better chance to be purchased. To allow a system to learn about users' online behavior and their preferred choices, we use all available data on products, consumer demographics and historical online transactions. Thus, an important phase in the process of enhancing the website with personalization capabilities is segmentation of customers and products into related groups according to customers' preferences in a way that leads to increased likelihood of the purchase.

One of the challenges in identifying clusters (groups) of consumers or products is knowledge on how 'close' or how 'far apart' consumers are to each other. Two consumers are 'close' when their dissimilarity or distance is small or, equivalently, if their similarity is large.

There are different proximity measures depending on data types – categorical, continuous or a mix of the two. Two typical measures for continuous data are

$$\textbf{Minkowski distance:} \quad \mu(x,y) = (\Sigma_i |x_i - y_i|^p)^{1/p}, \quad p \geq 1,$$

$$\textbf{Cosine similarity:} \quad \kappa(x,y) = \frac{< x, y >}{\|x\|\|y\|} = \frac{\Sigma_i x_i y_i}{(\Sigma_i x_i^2 \Sigma_i y_i^2)^{\frac{1}{2}}}$$

Performance of clustering algorithms depends critically on the used distance (proximity) metric so it is very important to make an appropriate selection. The choice depends on the shape of clusters which may be visually analyzed in lower dimensional spaces ($d \leq 3$).

Assume that the goal is to cluster two groups of consumers in R^2 product space. Each axis in R^2 represents consumers' preference for a product or a group of products. Figure 1 illustrates a case where one group of consumers has strong preference (with high variation) for Product 1, whereas the second group of consumers has strong preference for Product 2 but with lower variance.

Clustering based on Euclidean distance, in this case, may lead to misclassification of some members of high variability group as indicated by dashed line which wrongly assigns a part of group's members with strong preference in Product 1 to the group with preference in Product 2.

Described problem can be solved by using cosine similarity measure, but cosine

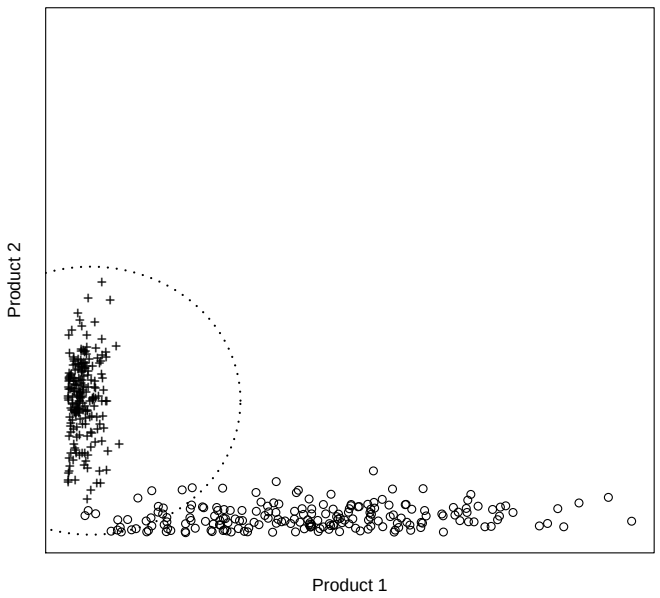


Figure 1: R^2 Euclidean

fails in cases when groups of consumers can not be separated by preferences in one of the products/group of products (see Figure 2). In Figure 2 consumers groups can be separated by preferences to Product 1, but not by preferences to Product 2. When applying cosine similarity measure consumers from "Product 2" get misclassified as the members of "Product 1" cluster.

2.1 Angular Distance Measure

We introduce Angular Distance Measure (ADM), which is a new proximity measure that can overcome limitations of Minkowski distance and cosine similarity in consumer preference groupings. ADM consists of a distance metric, a normalizing function, and an angular component (cosine function).

Angular Distance Measure is defined in the first orthant of R^d as:

$$\delta_{xy} = \frac{\mu(x,y)}{\varphi(x,y)}[1 - \cos(x,y)] \tag{1}$$

where $d > 1$, $x, y \in R^d_+$, $\mu(x,y)$ is a distance function such as Minkowski distance and $\varphi(x,y) > 0$ is a normalizing function, with examples given below.

- In case we measure closeness between two points: $\varphi(x,y) = (\|x\| + \|y\|)^{1/2}$.
- In case we measure distance from a cluster center to a point, $\varphi(x,y)$ can be:

1. Cluster's variation, $\sigma^2_{C_j} = \Sigma_{x \in C_j} \frac{\|x - c_j\|}{N_{C_j}}$, where c_j is a center of the cluster C_j, $x \in C_j$ and N_{C_j} is a number of observations/points in C_j.

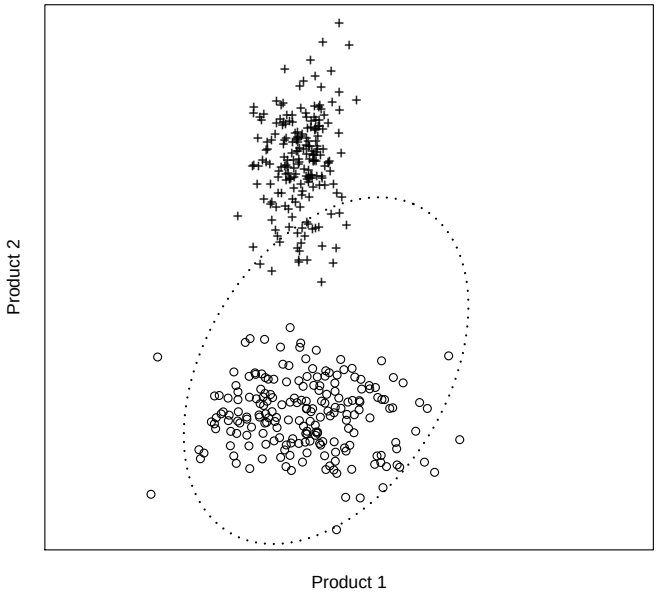


Figure 2: R^2 Cosine similarity

2. Cluster's radius, $\phi_{C_j} = \max_{x \in C_j} d(x, c_j)$.

Angular Distance measure is symmetric and non-negative by definition (note that $\cos(x, y) \in [0, 1]$) as the vectors x and y are in the first orthant. Also, ADM can be zero when $x \neq y$.

2.2 Application of ADM to grouping of consumer preferences

Applying a clustering algorithm with ADM as distance function to the data in the Figure 1 and Figure 2 results in proper segmentation (we do not show the graphs). In particular, (1) ADM correctly partitions customers with high variability of preference for Product 1 where Euclidean distance fails (Figure 1) and (2) groups correctly customers in case where cosine similarity fails (Figure 2).

The algorithm implementing segmentation of customers according to their purchase preferences can be divided in two parts: (1) construction of the new products/categories space; and (2) grouping of consumers into segments.

Algorithm first reduces initial dimension of product space by aggregating together similar products (categories) using ADM. For example, in the context of online consumer preferences, the products bought close in time may be regarded as similar. Similarity between each pair of products is measured by mapping consumers transactions to R^2 space where each dimension represents one of the products. "Majority" rule is applied to consumer transaction data to form two sets in R^2 and ADM is computed between centroids of the two sets. Products are aggregated together if they

are closer than some predefined threshold, τ_0, which can be set to the maximum of two sets' variances.

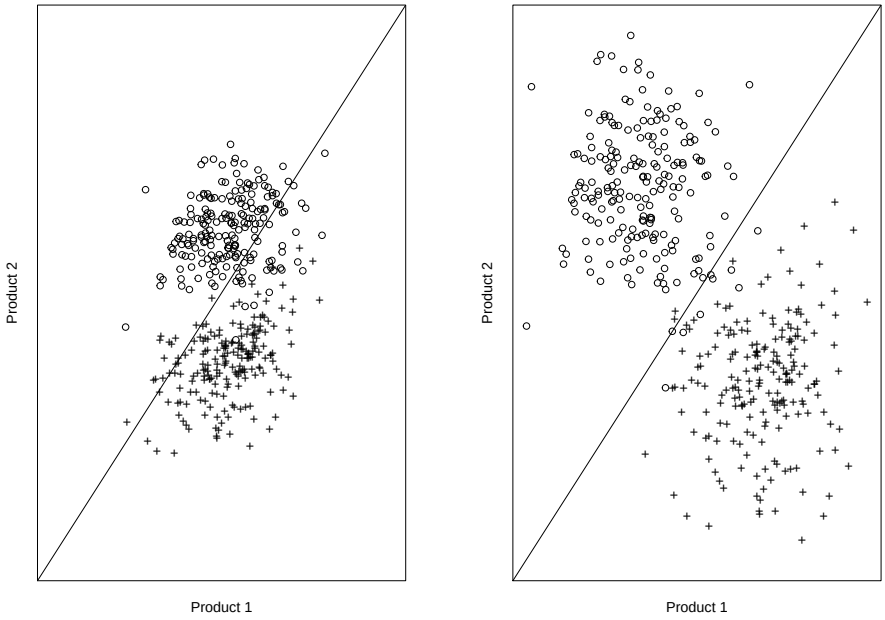


Figure 3: Two sets aggregation example

Aggregation is illustrated in Figure 3 where diagonal lines represent a "majority" rule which divides points in two clusters. Left panel in the figure represents the case when two products can be aggregated. Clusters (their centroids, in fact) are located close to each other. Sets on the right side panel cannot be aggregated as their centroids are far from each other and so the sets form separate clusters.

At the next step, the algorithm aggregates together products with smallest eligible values of ADM. The algorithm repeats the aggregation steps until all "eligible" products are aggregated. The algorithm is then repeated from the beginning, for smaller number of the dimensions, until desired number of dimensions is reached or a subsequent aggregation cannot be performed.

Below is a high level outline of the first part of the algorithm. Main elements of the input include (1) customers' online transactions for all products, (2) similarity threshold, τ_s, and (3) threshold, τ_o, for splitting product pair $2D$ space in two parts. Details of the algorithm and a more thorough analysis of its performance will be reported elsewhere.

Output of the first part of the algorithm is a set of N new product categories (dimensions). Each dimension represents consumers preference in particular set of products. Segmentation (partitioning) of consumers is performed in the second part of the algorithm using N as the number of clusters for K-means algorithm with ADM as the distance function.

Algorithm to construct new category space:

1. Let $x_n = (p_{ni}, p_{nj})$ be preferences of customer n for the pair of products (P_i, P_j)
2. Construct sets of customers $S_i = \{x_n : f(x_n, \tau_0) \geq 0\}$ and $S_j = \{x_n : f(x_n, \tau_0) < 0\}$
3. Compute ADM δ_{ij} between centers $c_i = T(S_i)$ and $c_j = T(S_j)$ of S_i and S_j
4. If $\delta_{ij} > \tau_s$ for all (P_i, P_j) then nothing to aggregate. Stop the algorithm.
5. Otherwise, find P_i and P_j corresponding to the smallest δ_{ij}
6. Aggregate pair (P_i, P_j) in one category and remove it from the set of pairs for which $\delta_{ij} \leq \tau_s$
7. Repeat step 6 until set $\{(P_i, P_j) : \delta_{ij} \leq \tau_s\}$ is empty
8. Use aggregated products as new dimensions and repeat the algorithm until desired number of dimensions has been reached.

2.3 Application results

The algorithm was applied to several eBay data sets including one for a retailer's baby products. Historical consumer transactional data and product topology were used as inputs to the algorithm. Results of the algorithm were used in a number of display campaigns targeting consumers on the third party webpages.

As an example, we describe results on a dataset with initial product category space of 18 dimensions represented by the following categories:
/ Infant care / Layette clothes / Diapers & wipes / Baby accessories / Infant bedding / Newborn / Gear & home / Wooden furniture / Bulk juvenile products / Juvenile room decor / Consumables / Nursery & gifts / Infant girls, boys / Food formula / Sleepware / Imaginarium / Kids furniture / Giftware baby /.

Application of the ADM-based algorithm resulted in a new six-dimensional category space. The new dimensions are presented in Table 1. Partitioning of consumers (about 270K) was performed using the newly obtained category space of six dimensions and by setting the number of clusters to six. Numerical results of partitioning are presented in Table 2. We computed 6×6 matrix (not shown) with the centers of consumers' clusters for six dimensions in which a diagonal element represents a cluster's projection to the "preferred" dimension.

For instance: Dimension 1 is the 'preferred' dimension of consumers assigned to the Cluster 1 and Consumers from Cluster 1 made on average 2.56 transactions in category represented by categories, Diapers and wipes, Food formula, Imaginarium, along with less than 0.4 transaction in other categories. Similarly for other cluster-dimension pairs. Results of the algorithm were used for re-targeting of the consumers and for collaborative filtering type of targeting.

Comparison of the performance of the ADM based algorithm and K-means clustering (with Euclidean distance) is summarized in Table 3 Metrics used for performance measurements include: Click-through rate, Post impression conversion, Total conversion, Average order value (AOV) and Return on investment (ROI). ADM based algorithm on average outperformed the K-means clustering on all metrics. Moreover,

the differences between ADM and K-means on the metrics Post impression conversion, Total conversion and Return on investment are also statistically significant.

Table 1: New category space dimensions

Dim.	Categories included
1	*/Diapers & wipes /Food formula / Imaginarium /*
2	*/Layette clothes/*
3	*/Infant care/ Consumables / Baby accessories / Giftware baby /*
4	*/Infant bedding/ Kids furniture / Nursery & gifts /*
5	*/Newborn/ Infant girls, boys / Sleepware /*
6	*/Gear and home/ Wooden furniture/ Bulk juvenile products/ Juvenile room decor/*

Table 2: Counts of consumers per cluster

Cluster	1	2	3	4	5	6
Consumers	6398	32422	99805	62141	6670	168294

Table 3: ADM-based algorithm vs K-means with Euclidean distance

	Click-through rate (%)	Post impression conversion (%)	Total conversion (%)	AOV ($)	ROI ($)
ADM	**0.0424**	**0.2595**	**0.249**	**93.09**	**49.93**
K-means	0.0393	0.1347	0.130	78.92	26.15

Conclusions

The first section provides a brief survey of practical forecasting approaches and methods when faced with a problem of making large number of forecasts in industrial setting. We prefer combining forecasts from multiple methods or models, over searching for a single "best" model. In particular, we advocate use of statistical time-series models capable of producing forecasts for hierarchies of aggregated data and multiple time scales. Careful evaluation of models' predictive performance, using adequate accuracy metrics, is important in model selection process and must be done on proper holdout sets. ML techniques still have room for improvement of forecasting perfomance for time series, yet some of ML algorithms (such as RF-s) are among our favorite tools for classification tasks.

In the second section we illustrate an application of segmentation for e-commerce personalization, where products and customers are clustered in a way that desired merchandise can be offered to customers without explicit querying. Segmenting the product-customer space into associated groups depends heavily on the proximity measures. We described ADM, a novel such measure, which outperforms its Euclidean and cosine counterparts. Results were illustrated on toy examples and real data.

References

[1] Athanasopoulos G, Hyndman R, Kourentzes N, Petropoulos F, (2017) Forecasting with temporal hierarchies. European J of Operat. Research, Vol 262/1, 60-74.

[2] Bates JM, Granger C (1969) The Combination of Forecasts. Operations Research Quarterly 20(4):451-468.

[3] Cawley GC, Talbot NC (2010): On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation. J of ML Research, 11 2079-2107.

[4] Croston, J. D. (1972). Forecasting and stock control for intermittent demands. Operational Research Quarterly, 23(3), 289–303.

[5] Deng L. (2014) A tutorial survey of architectures, algorithms, and applications for deep learning. APSIPA Trans on Signal and Information Processing. 2014; 3.

[6] Green KC, Armstrong JS. (2015) Simple versus complex forecasting, The evidence. Journal of Business Research. 2015;68(8):1678–1685.

[7] Hyndman R. (2006) Another look at forecast-accuracy metrics for intermittent demand, Foresight, Intl J of Applied Forecasting, 2006, 4, 43-46.

[8] Hyndman R, Athanasopoulos, G. (2018) Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2

[9] Krnjajić M, Draper D, (2014) Bayesian model comparison: Log scores and *DIC*. Statistics & Probability Letters (2014) 88(1):9–14

[10] Leskovec J, Rajaraman A, Ullman J, (2014) Mining of Massive Datasets, book available online http://infolab.stanford.edu/%7Eullman/mmds/book.pdf

[11] Makridakis S, Spiliotis E, Assimakopoulos V, (2018) Statistical and Machine Learning forecasting methods: Concerns and ways forward. PLoS ONE 13(3)

[12] Petropoulos F, Kourentzes N. (2015) Forecast combinations for intermittent demand. Journal Operational Research Society Vol 66, 2015 - Issue 6

[13] Petropoulos F, Kourentzes N, Nikolopoulos K, (2016) Another look at estimators for intermittent demand, Intl Journal of Production Economics 181, 154-161.

[14] R Core Team (2017). R: A language and environment for stat. computing. R Foundation for Stat. Computing, Vienna, Austria. https://www.R-project.org/.

[15] Tofalis C, (2015) A better measure of relative prediction accuracy for model selection. Journal of Operation Research Societ, (2015), 66.

[16] Wickramasuriya S, Athanasopoulos G, Hyndman RJ. (2019) Optimal forecast reconciliation for hierarchical and grouped time series. JASA , 114(526), 804-819.

Effect of sampling jitter in devices for discrete signal processing

Aleksey G. Vostretsov and Vasily N. Vasyukov
Novosibirsk State Technical University, Novosibirsk, Russia
e-mail: vostreczov@corp.nstu.ru, vasyukov@corp.nstu.ru

Abstract

Real discretization is always accompanied by so-called sampling jitter [2] that decreases efficiency of the algorithms constructed without considering it. To reduce its effect, high stability driving generators are recommended to be used. The quartz generators used in practice have relative instability of $10^{-8}...10^{-11}$. However, as it will be shown below, even in this case the sampling jitter can significantly affect characteristics of systems for discrete signal processing. This problem, nevertheless, is not well represented in publications. The work [1] should be mentioned. The paper [1] deals with effect of jitter on correlative and spectral characteristics of the sequence of samples of random and deterministic signals. We offer the mathematical model describing jitter effect in systems with high stability driving generators. We analyse processes at the output of the sampling gate when observing additive mixture of signal and noise; the characteristics of the processes at the output of sampling gate are investigated taking into account sampling jitter effect.

Keywords: Sampling, jitter, noise, instability, discrete signal processing.

1 Model of sampling jitter in systems with high stability driving generators

Let us consider the generalized structure diagram of a system for discrete signal processing, fig. 1, a. The main elements are input amplifier 1, filter 2, sample-and-hold circuit (SHC) 3, clock-pulse generator 4, and device for processing signal samples 5. Input amplifier 1 and filter 2 form the linear part of the system, determine required dynamic range and prevent overlapping of spectra during input signal discretization. SHC and the generator form the discretization device.

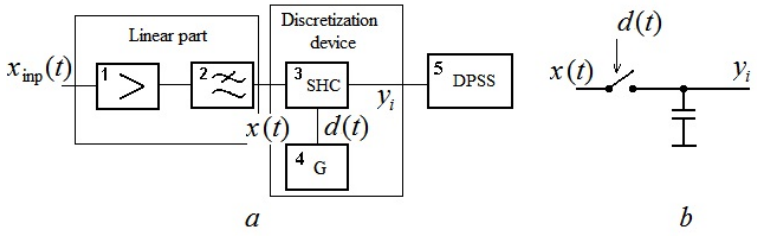


Figure 1

For development of sampling jitter model we assume that SHC consists of an ideal switch controlled by the clock-pulse generator 4 and a capacitor for storing the signal level over a period required for analog-to-digital conversion, fig. 1, b.

Let us denote the signal at the input of the system as $x_{inp}(t)$, at the output of the linear part of the system as $x(t)$, the value of the i-th sample at the SHC output as $y_i, y_i = x(t_i)$, here t_i is the time point of taking the i-th sample. The switch connects the capacitor to the output of the filter 2 when the sync signal $d(t)$ coming from the generator exceeds the threshold level h. The sync signal is the sum of the highly stable oscillation $S(t)$, produced by the generator, and the stationary noise $v(t)$ always being present in synchronizing circuit. Therefore the time point t_d when the threshold level h is crossed by the mixture $d(t) = S(t)+v(t)$ is biased with respect to the time point t_s when this level is crossed by the oscillation $S(t)$, by the random variable $\xi = t_d - t_s$, fig.2.

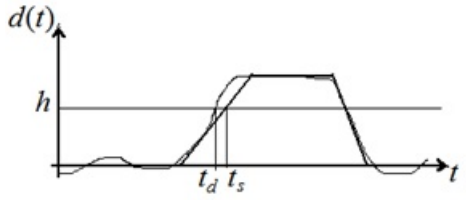


Figure 2

Let instability of the generator be negligibly low, so that we do not consider it. Then the jitter is determined only by the random component of the time points when the sync signal crosses the level h, the component being caused by noise in synchronizing circuits.

Assuming that the oscillation $S(t)$ has the form $S(t) = U_s s(t)$, where $s(t)$ is the periodic signal normalized in amplitude, $v(t)$ is the differentiable stationary Gaussian process with zero mean value and the variance σ_v^2; the signal-to-noise ratio (SNR) in synchronizing circuits is $q_v = U_s/\sigma_v$. Then for large values of the parameter q_v the probability distribution of the random variable ξ is represented by the Gaussian with the zero mean and the root-mean-square deviation

$$\sigma_\xi = 1/[q_v|s'(t_s)|], \tag{1}$$

where $s'(t_s)$ is the derivative of the normalized in amplitude oscillation $s(t)$ at the time point when the oscillation crosses the level h bottom-up. In particular, if $h = 0, s(t) = \sin(\omega t)$, where $\omega = 2\pi/T_d$ is the sampling frequency, then $\sigma_\xi = 1/[q_v\omega]$. The relative root-mean-square bias of the sampling instant $\delta_\xi = \sigma_\xi/T_d$ is determined in this case by the expression $\delta_\xi = 1/[2\pi q_v]$, and for the signal-to-noise ratio in the synchronizing circuits of $q_v = 40$ dB has the value of 0.0016 that is significantly (by several orders) higher than the relative frequency instability of the quartz generator ($10^{-8}...10^{-11}$). Thus, the sampling jitter has noticeable effect that cannot be neglected in calculation of efficiency even in systems with highly stable clock-pulse generators.

2 Mathematical model and equivalent circuit of discretization device

Let the process $x(t)$ observed at the output of the linear part of the system be the sum of the useful signal $u(t)$ and the stationary zero mean noise $\eta(t)$:

$$x(t) = u(t) + \eta(t).$$

The useful signal $u(t) = U_m u(t, \lambda)$ can be either quasi-deterministic or random. In the former U_m is the signal amplitude, $u(t, \lambda)$ is the signal normalized in amplitude, $\lambda = \{\lambda_1, ..., \lambda_M\}$ is the vector of the signal parameters; in the latter U_m is the root-mean-square value of the signal, $u(t, \lambda)$ is the realization of the stochastic process normalized in power, $\lambda = \{\lambda_1, ..., \lambda_M\}$ is the vector of the process parameters. The components of the parameters vector λ can be partially or completely unknown. The mixture $x(t)$ is characterized by the signal-to-noise ratio $q = U_m/\sigma_\eta$, where σ_η is the root-mean-square deviation of the noise.

At the output of the discretization device the sequence of N samples $y_i = x(t_i)$ taken at the time points $t_i, i = 1, ..., N$, is formed; each sample represents the sum of the useful signal sample $u(t_i) = U_m u(t_i, \lambda)$ and the noise component sample $\eta(t_i)$ at the output of the linear part of the system. For realization of algorithms for discrete signal processing it is usually assumed that the samples are taken at the time points $t_{i0} = iT_d$ (T_d is the sampling interval), but actually they are taken at the time points $t_i = t_{i0} + \xi_i$, $i = 1, ..., N$, where ξ_i is the random bias of the i-th sample because of jitter.

Let us assume that the random variables ξ_i, $i = 1, ..., N$ form stationary zero mean sequence with small root-mean-square deviation σ_ξ, and that the functions $u(t)$ and $\eta(t)$ are doubly differentiable. Then, using three first terms of Taylor series expansion for representation of the process $x(t)$ in the neighbourhood of the point t_{i0} due to smallness of σ_ξ, we can write for the i-th sample at the output of discretization device

$$\begin{aligned} y_i &= x(t_i) = u(t_{i0} + \xi_i) + \eta(t_{i0} + \xi_i) \approx \\ &\approx u(t_{i0}) + \eta(t_{i0}) + \frac{1}{2}\Big[\frac{d^2u(t_{i0})}{dt^2} + \frac{d^2\eta(t_{i0})}{dt^2}\Big]\overline{\xi_i^2} + \Big[\frac{du(t_{i0})}{dt} + \frac{d\eta(t_{i0})}{dt}\Big]\xi_i + \\ &+ \frac{1}{2}\Big[\frac{d^2u(t_{i0})}{dt^2} + \frac{d^2\eta(t_{i0})}{dt^2}\Big](\xi_i^2 - \overline{\xi_i^2}). \end{aligned} \tag{2}$$

The first three terms in the right hand side of expression (2) describe frequency distortion of the input oscillation caused by sampling jitter, the rest two terms describe appearance of additional noise.

The frequency distortions of the input process $x(t)$ caused by sampling jitter are equivalent, as seen from expression (2), to the effect of the filter with the complex frequency response $K(j\omega) = 1 - \frac{1}{2}\omega^2\sigma_\xi^2$, where $\omega = 2\pi f$ is the angular frequency; $\sigma_\xi^2 = \overline{\xi_i^2}$ is the variance of samples of the random sequence ξ_i that is constant owing to its stationarity. The last is valid if the noise $v(t)$ is stationary and sync signal is

strictly periodic. The additional noise in samples of the process is represented by the random variables

$$\varsigma_i = \Big[\frac{\partial u(t_{i0})}{dt} + \frac{\partial \eta(t_{i0})}{\partial t}\Big]\xi_i + \frac{1}{2}\Big[\frac{\partial^2 u(t_{i0})}{\partial t^2} + \frac{\partial^2 \eta(t_{i0})}{\partial t^2}\Big](\xi_i^2 - \overline{\xi_i^2}) \approx \Big[\frac{\partial u(t_{i0})}{\partial t} + \frac{\partial \eta(t_{i0})}{\partial t}\Big]\xi_i. \tag{3}$$

It is obvious that ς_i and ς_j for $i \neq j$ are uncorrelated with each other and with the noise samples $\eta(t_i)$. When the useful signal is absent the variables ς_i, $i = 1...N$ have the same variance $\sigma_\varsigma \approx \sigma_{\eta'}^2\sigma_\xi^2$($\sigma_{\eta'}^2$ denotes the variance of the derivative $\eta'(t)$ of the process $\eta(t)$). Therefore, they can be considered as the samples of stationary noise having in the frequency band of the width $\Delta F = \frac{1}{2T_d}$ constant power spectral density

$$N_\varsigma = \sigma_\varsigma^2/\Delta F = 2\sigma_\varsigma^2 T_d. \tag{4}$$

In the case when linear part of the system has $\prod$-shaped amplitude frequency response with the boundary frequencies F_H and F_B, after substituting in (4) the variance σ_ς^2 expressed through the parameters of the input noise $\eta(t)$ and the jitter ξ_i, we obtain

$$N_\varsigma \approx 8\pi^2\sigma_\varsigma^2 T_d \int_{F_H}^{F_B} f^2 N_\eta(f)df, \tag{5}$$

where $N_\eta(f)$ is the power spectral density of the initial noise.

Appearance of deterministic signal increases the variance of additional noise caused by sampling jitter by $\sigma_{ui}^2 = \Big(\frac{\partial u(t_{i0})}{\partial t}\Big)^2\sigma_\xi^2$, and the total variance of the i-th sample $(i = 1, ..., N)$ becomes

$$\sigma_{\varsigma i}^2 = \Big(\frac{du(t_{i0})}{dt}\Big)^2\sigma_\xi^2 + \sigma_{\eta'}^2\sigma_\xi^2 = \Big[\Big(\frac{du(t_{i0})}{dt}\Big)^2 + 4\pi^2\int_{F_H}^{F_B} f^2 N_\eta(f)df\Big]\sigma_\xi^2. \tag{6}$$

As it follows from (6) the existence of the deterministic signal at the input leads to appearance at the output of additional nonstationary component of noise with the variance depending on the derivative of the signal at the time points when samples are taken.

For the doubly differentiable stationary random process $u(t)$ the total variance of the i-th sample is determined by expression

$$\sigma_{\varsigma_i}^2 = 4\pi^2\sigma_\xi^2\int_{F_H}^{F_B} f^2[N_\eta(f) + N_u(f)]df, \tag{7}$$

where $N_u(f)$ is the power spectral density of the signal at the output of linear part of the system.

According to mentioned above, fig. 3 represents the equivalent circuit of the discretization device taking into account jitter effect.

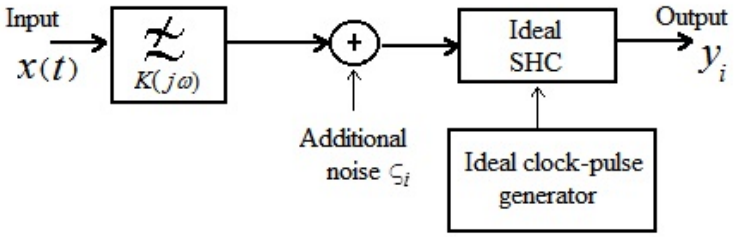


Figure 3

3 Signal-to-noise ratio at the output of discretization device

For characterizing samples of the observed process we will use the signal-to-noise ratio $q^2_{\sum} = \sigma^2_{\sum_s}/\sigma^2_{\sum_n}$, where for a random input signal $\sigma^2_{\sum_s} = \int\limits_{F_H}^{F_B} N_u(f)|K(j2\pi f)|^2 df$ is the variance of the samples sequence of the signal component of the process $u(t)$ at the output of discretization device ($N_u(f)$ is the power spectral density of the signal at the output of the linear channel of the system, the factor $|K(j2\pi f)|^2$ takes into account the effect of suppression of the signal high frequency components owing to sampling jitter); $\sigma^2_{\sum_n} = \int\limits_{F_H}^{F_B} N_\eta(f)|K(j2\pi f)|^2 df + \sigma^2_{\varsigma_i}$ is the variance of the samples sequence of the noise component of the process at the output of the discretization device. For a deterministic periodic signal $\sigma^2_{\sum_s}$ represents the average power of the signal, and $N_u(f)$ – its power spectral density consisting of the weighted δ-functions corresponding to harmonic components. Taking into account $K(j2\pi f) = 1 - \frac{1}{2}(2\pi f)^2\sigma^2_\xi$, and neglecting second order of smallness values we obtain:

$$q^2_{\sum} \approx \frac{\int\limits_{F_H}^{F_B} N_u(f)df - 4\pi^2\sigma^2_\xi \int\limits_{F_H}^{F_B} f^2 N_u(f)df}{\int\limits_{F_H}^{F_B} N_\eta(f)df + 4\pi^2\sigma^2_\xi \int\limits_{F_H}^{F_B} f^2 N_u(f)df} = \frac{q^2 - \frac{4\pi^2\sigma^2_\xi \int\limits_{F_H}^{F_B} f^2 N_u(f)df}{\int\limits_{F_H}^{F_B} N_\eta(f)df}}{1 + \frac{4\pi^2\sigma^2_\xi \int\limits_{F_H}^{F_B} f^2 N_u(f)df}{\int\limits_{F_H}^{F_B} N_\eta(f)df}},$$

where q^2 is the SNR at the output of the linear channel of the system. It is obvious that if the variance of jitter σ^2_ξ increases the SNR at the output of the discretization device decreases.

Let us consider the situation when both the useful signal and the signal in the synchronizing circuits are sinusoidal. It is possible to show that in this case $q^2_{\sum} = q^2\big(1 - \frac{T_d^2 f_0^2}{q_v^2}\big)/\big(1 + \frac{q^2 T_d^2 f_0^2}{q_v^2}\big)$, where f_0 is the sync signal frequency.

It is interesting to analyse the behaviour of the resulting SNR for various q^2 and q^2_v. It is apparent that for the weak jitter ($q^2_v \longrightarrow \infty$) the resulting signal-to-noise

ratio tends to $q_{\sum}^2 = q^2/\left(1+\frac{q^2 T_d^2 f_0^2}{q_v^2}\right)$. Then two different cases can be considered. For a relatively weak input signal $\left(q << \frac{q_v}{T_d f_0}\right)$ the resulting SNR is determined by the input ratio $q_{\sum}^2 \longrightarrow q^2$. For a high SNR at the input $\left(q >> \frac{q_v}{T_d f_0}\right)$, that is the case for high precision measurements, the resulting SNR is entirely determined by the SNR in the synchronizing circuits and tends to $q_{\sum}^2 = q_v^2/(T_d^2 f_0^2)$. In addition, the SNR at the output of the discretization device decreases with increase of the signal frequency owing to suppression of high frequency components of the observed process by jitter.

Conclusion

We offer a simple linear model for description of influence of sampling jitter on the signal and signal-to-noise ratio in the system of digital signal processing. Within the bounds of this model, we show that sampling jitter causes frequency distortion of the signal and appearance of additional noise.

References

[1] Balakrishnan A.V. (1962). On the Problem of Time Jitter in Sampling. *IRE Transactions on Information Theory*. Vol. **IT-8**, pp. 226-236.

[2] Otnes R.K. and Enochson L. (1978). Applied Time Series Analysis. New York: Wiley

An omega-square statistics for analysis of correspondence of small texts to the Zipf—Mandelbrot law

Natalia Zakrevskaya[1] and Artyom Kovalevskii[1,2]
[1] *Novosibirsk State Technical University, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
e-mail: natali_erlagol@mail.ru, pandorra@ngs.ru

Abstract

The elementary probabilistic model of text assumes that words of a text appear independently of each other in a random way. The model is determined by the probabilities each word to take its meaning. These probabilities satisfy the Zipf—Mandelbrot law. We develop and implement a statistics of omega-square type for analysis of correspondence of texts to this elementary probabilistic model. We present results of the analysis of Shakespeare's sonnets.

Keywords: *Zipf's law, statistical test, weak convergence, text analysis.*

Introduction

It has long been observed that any author uses more and more different words as the text is written. Even considering of collected works with hundreds of thousands and millions of words shows that the emergence of new words never stops. At the same time, the growth rate of the number of different words decreases as the length of the text increases. In our study, we call word forms as words, that is, we consider different forms of a word as different words.

A simple probabilistic model for the number of different words in a text is the number of non-empty urns in an infinite urn scheme. The choice of each next word of the text is associated with a random choice of an urn for the next ball, and the number of urns is infinite. The urn is chosen randomly for each ball, independently of the others. The probability p_i of the choice of the urn i is the same for all balls, $i \geq 1$.

Zipf (1936) investigated the frequency of the appearance of words in the text. He proposed a power law of decrease of frequencies, that is, frequency is proportional to the rank of the word with some negative exponent. Mandelbrot (1965) modified Zipf's law and gave a number interpretations of it. We will interpret frequencies as probabilities of the appearance of words. In accordance with Zipf—Mandelbrot law, the word with rank i has probability

$$p_i = c(i-\delta)^{-\alpha}, \quad i \geq 1. \tag{1}$$

Here $\alpha > 1$ is the Zipf exponent, $\delta > -1$ is the Mandelbrot shift, C is normalizing constant. If $\delta = 0$ then we have a classical Zipf's law, in this case $c = \zeta^{-1}(\alpha)$, and ζ is Riemann zeta function.

Let R_i be the number of different words among the first i words of the text under consideration, and n be the number of all words in the text. Sequence $R_1, R_2, \ldots, R_n$ does not decrease, $R_1 = 1$. We let $R_0 = 0$.

Bahadur (1960) substantiated the law of growth in the number of different words for a simple probabilistic model of text. He noted that the expectation of the number of different words is calculated by

$$\mathbf{E}R_n = \sum_{i=1}^{\infty}(1 - (1 - p_i)^n) \tag{2}$$

and has asymptotics

$$\mathbf{E}R_n \sim \Gamma(1 - \theta)c^\theta n^\theta. \tag{3}$$

Here $\theta = \alpha^{-1}$, $\theta \in (0, 1)$.

He proved the Law of Large Numbers, that is, convergence $R_n/\mathbf{E}R_n \to 1$ in probability.

Karlin (1967) proved the Strong Law of Large Numbers and the Central Limit Theorem for R_n. The theory of the limiting behavior of a number of different words and similar statistics within the framework of this elementary models developed by Dutko (1989), Zakrevskaya and Kovalevskii (2001), Barbour and Gnedin (2009), Barbour (2009), Ohannessian and Dahleh (2012), Chebunin (2014), Muratov and Zuyev (2016), Chebunin and Kovalevskii (2016, 2018, 2019), Ben-Hamou, Boucheron and Ohannessian (2017).

Estimation of the Zipf parameter based on the number of different words is studied in several papers. Zakrevskaya and Kovalevskii (2001) proposed an substitution estimate and proved its consistence. Ohannessian and Dahleh (2012) proposed an estimate as the ratio of the number of different words to the number of words that present once in the text, and proved its consistence. Chebunin and Kovalevskii (2018) introduced two classes of estimates that include estimates by Zakrevskaya and Kovalevskii, Ohannessian and Dahleh, and proved their strong consistency and asymptotic normality.

The procedure for testing the hypothesis of correspondence of a text to an elementary probabilistic model based on the sequence of numbers of different words $R_1, \ldots, R_n$ proposed by Chebunin and Kovalevsky (2019).

Our paper discusses approaches of its practical application to the texts, as well as the results of the application to the analysis of Shakespeare's sonnets.

Note that the power-law character of the growth in the number of different words is noted in linguistics, it is associated with the names of Herdan (1960) and Heaps (1978). Petersen et al. (2012) revealed deviations from this law for very large texts.

The remainder of the article is constructed as follows: we discuss the selection of words in a text in English and counting the sequence of quantities of different words in Section 1, we analyze and modify the parameter estimation procedure and calculating the estimate of the expectation of the number of different words in Section 2, we develop calculation of a statistics of omega-square type in Section 3.The results of applying the developed statistics to Shakespeare sonnets are in Section 4.

1 Sequence of numbers of different words

To count the number of different words, we need to learn how to extract words from the text. A word is technically defined as a sequence of characters between two spaces. But it is necessary to exclude punctuation marks, they are not parts of a word. There are, however, two exceptions: hyphen between letters is considered as a part of the word. Also an apostrophe between the letters, and before the word and after the word is a part of the word (but we need to distinguish the apostrophe from single quotes).

Words are considered the same if they consist of the same characters in the same order. Only difference allowed for the first character: if the ascii codes of the first characters differ by 32, words are considered the same (because the difference between ascii-codes of uppercase and lowercase English letters by 32). The program forms a vocabulary of text and record the number of repetitions of any word. A word is recorded in the vocabulary with the first letter (upper or lower case) with which it first met in the text under study.

On the basis of the vocabulary of the text, a sequence of numbers of different words is calculated.

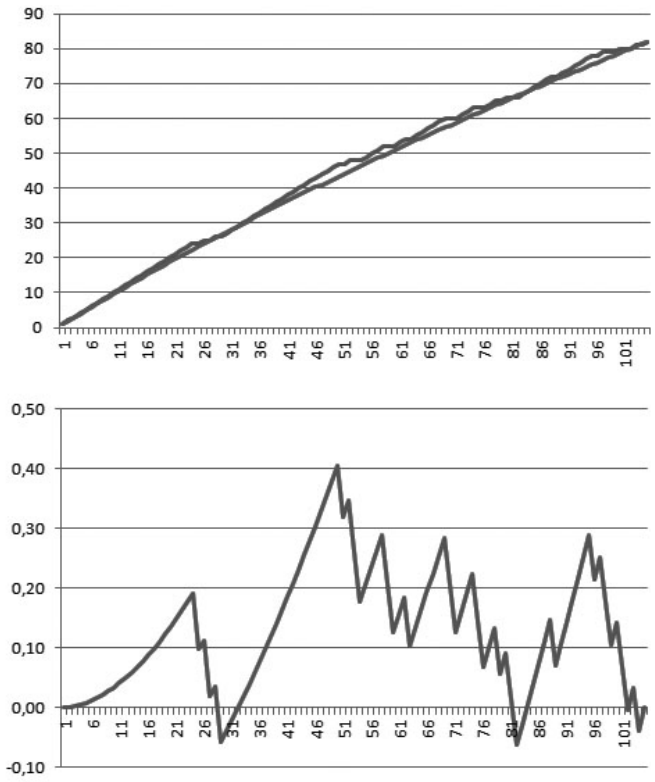


Figure 1: Sonnet 1

2 Estimates

So, the researcher has the values $R_1, \ldots, R_n$. It is assumed that the two-parameter model (1) is true. We estimate the parameter $\theta = 1/\alpha$ by a modification of the estimate $\widehat{\theta} = \ln(R_n/R_{[n/2]})/\ln 2$, proposed by Chebunin and Kovalevskii (2019). The strong consistency and asymptotic normality of this estimate are proved in the cited paper. The disadvantage of this estimate is that $[n/2]$ is shifted to the left relative to the middle of the interval from 0 to n, if n is odd. Therefore, a correction of this estimate is proposed: if n is even then $\widetilde{\theta} = \widehat{\theta}$; if n is odd then

$$\widetilde{\theta} = \frac{\ln(2R_n) - \ln(R_{[n/2]} + R_{[n/2+1]})}{\ln 2}.$$

This estimate is also strongly consistent and asymptotically normal since $R_{[n/2+1]}$ differs from $R_{[n/2]}$ no more than by 1, and $R_n \to \infty$ almost surely.

The estimation of the shift parameter δ is carried out together with the calculation of the fitted values $\widetilde{R}_1, \ldots, \widetilde{R}_n$. We caculate fitted values by (2) with substitution of estimates $\widetilde{\alpha} = \widetilde{\theta}^{-1}$, $\widetilde{\delta}$ instead of parameters α and δ in (1). We change the sum to infinity in (1) by sum to $M = 10^6$.

As $\delta > -1$, we take $\delta_- = -0.9$, $\delta_+ = 100$, and find $\widetilde{\delta}$ by dichotomy method as the root of the equation $\widetilde{R_n} = R_n$. The root is single due to monotonicity of $\widetilde{R_n}$ as function of $\widetilde{\delta}$.

3 Statistical test

To test the hypothesis of correspondence of a text to the Zipf—Mandelbrot model, we form an empirical bridge just as it was done in the paper of Kovalevskii and Shatalin (2015). Let us introduce the process $\widetilde{Z}_n$:

$$\widetilde{Z}_n(k/n) = \left(R_k - \widetilde{R}_k\right)/\sqrt{R_n},$$

$0 \le k \le n$. Let for $0 \le t \le 1/n$ and $0 \le k \le n-1$

$$\widetilde{Z}_n\left(\frac{k}{n} + t\right) = \widetilde{Z}_n(k/n) + nt\left(\widetilde{Z}_n((k+1)/n) - \widetilde{Z}_n(k/n)\right).$$

We have from (3) and SLLN

$$\left(\widetilde{R}_k - (k/n)^{\widetilde{\theta}} R_n\right)/\sqrt{R_n} \to 0$$

almost surely uniformely on c in any compact in $(-1, \infty)$ and on k/n in any compact in $(0, 1]$.

So $\widetilde{Z}_n - \widehat{Z}_n \to 0$ almost surely in uniform metrics uniformely on c in any compact in $(-1, \infty)$, process $\widehat{Z}_n$ is defined in Kovalevskii and Chebunin (2019). There is the Central Limit Theorem for this process, so the limiting distribution depends on θ only.

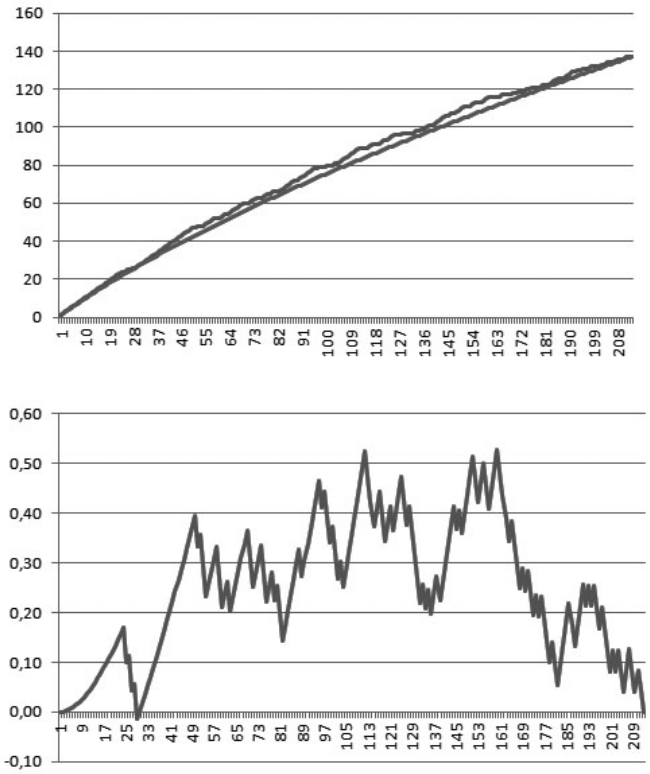


Figure 2: Sonnets 1 and 6

Let $\widetilde{W}_n^2 = \int\limits_0^1 \left(\widetilde{Z}_n^0(t)\right)^2 dt$. It is equal to

$$\widetilde{W}_n^2 = \frac{1}{3n} \sum_{k=1}^{n-1} \widetilde{Z}_n^0\left(\frac{k}{n}\right)\left(2\widetilde{Z}_n^0\left(\frac{k}{n}\right) + \widetilde{Z}_n^0\left(\frac{k+1}{n}\right)\right). \tag{4}$$

Two approaches are possible for evaluating of p-values. One of them is based on the work of Kovalevskii and Chebunin (2019) and requires finding eigenvalues ??and eigenfunctions for an known kernel. The other is based on modeling of sequences of numbers $X_1, \ldots, X_n$ in accordance with the law (1), in which estimates are substituted instead of the unknown parameters. Then we do with these numbers the same as with the words of the source text. Such modeling should be done about $2 \cdot 10^6$ times, because 10^6 simulations are few for a fairly confident calculation of the p-value with an accuracy of 0.001 (see, for example, Philonenko and Postovalov 2019). The p-value estimate is equal to the ratio of the number of cases when the simulated value (4) exceeded the calculated value in the text, to the total number of experiments.

4 Analysis of Shakespeare sonnets

We analyze 6 Shakespeare's sonnets. To give an example of heterogeneous text, we analyze sonnets 1 and 6 together. The results are shown in Table 1 and in Fig. 1 and 2.

Table 1: Shakespeare's sonnets

sonnet	1	2	3	4	5	6	1 & 6
words	105	114	115	101	104	108	213
θ	0.773	0.782	0.796	0.824	0.933	0.666	0.714
different words	82	86	79	77	84	73	137
δ	7.46	5.82	2.65	3.18	0.656	9.89	10.33
$\widetilde{W}_n^2$	0.027	0.049	0.0066	0.034	0.0074	0.021	0.080

Acknowledgements

The research was supported by RFBR grant 17-01-00683. The authors like to thank Mikhail Chebunin for helpful and constructive comments and suggestions.

The authors like to thank Dmitrii Aldagarov for help in programming.

References

[1] Bahadur, R. R., 1960. On the number of distinct values in a large sample from an infinite discrete distribution. *Proceedings of the National Institute of Sciences of India*, 26A, Supp. II, 67–75.

[2] Barbour, A. D., 2009. Univariate approximations in the infinite occupancy scheme. *Alea* 6, 415–433.

[3] Barbour, A. D., Gnedin, A. V., 2009. Small counts in the infinite occupancy scheme. *Electronic Journal of Probability*, Vol. 14, Paper no. 13, 365–384.

[4] Ben-Hamou, A., Boucheron, S., Ohannessian, M. I., 2017. Concentration inequalities in the infinite urn scheme for occupancy counts and the missing mass, with applications. *Bernoulli*, V. 23, 249–287.

[5] Chebunin, M. G., 2014. Estimation of parameters of probabilistic models which is based on the number of different elements in a sample. *Sib. Zh. Ind. Mat.*, 17:3, 135–147 (in Russian).

[6] Chebunin, M., Kovalevskii, A., 2016. Functional central limit theorems for certain statistics in an infinite urn scheme. *Statistics and Probability Letters*, V. 119, 344–348.

[7] Chebunin, M., Kovalevskii, A., 2018. Asymptotically normal estimators for Zipf's law. *Sankhya A*, 1–10.

[8] Chebunin, M., Kovalevskii, A., 2019. A statistical test for the Zipf's law by deviations from the Heaps' law. Submitted to *Stat. Inference for Stoc. Proc.*

[9] Dutko, M., 1989. Central limit theorems for infinite urn models, *Ann. Probab.* 17, 1255–1263.

[10] Heaps, H. S., 1978. *Information Retrieval: Computational and Theoretical Aspects*, Academic Press.

[11] Herdan, G., 1960. *Type-token mathematics*, The Hague: Mouton.

[12] Karlin, S., 1967. Central Limit Theorems for Certain Infinite Urn Schemes. *Jounal of Mathematics and Mechanics*, Vol. 17, No. 4, 373–401.

[13] Kovalevskii, A. P., Shatalin, E. V., 2015. Asymptotics of sums of residuals of one-parameter linear regression on order statistics, *Theory of probability and its applications*, Vol. 59, No. 3, 375–387.

[14] Mandelbrot, B., 1965. Information Theory and Psycholinguistics. In B.B. Wolman and E. Nagel. *Scientific psychology*. Basic Books.

[15] Muratov, A., and Zuyev, S., 2016. Bit flipping and time to recover, *J. Appl. Probab.*, V. 53, 650–666.

[16] Ohannessian, M. I., Dahleh, M. A., 2012. Rare probability estimation under regularly varying heavy tails, *Proceedings of the 25th Annual Conference on Learning Theory*, PMLR 23:21.1–21.24.

[17] Petersen, A. M., Tenenbaum, J. N., Havlin, S., Stanley, H. E., Perc, M., 2012. Languages cool as they expand: Allometric scaling and the decreasing need for new words. *Scientific Reports* 2, Article No. 943.

[18] Philonenko, P., Postovalov, S., 2019. The new robust two-sample test for randomly right-censored data. *Journal of Statistical Computation and Simulation*, Volume 89, Issue 8, 1357–1375.

[19] Zakrevskaya, N. S., Kovalevskii, A. P., 2001. One-parameter probabilistic models of text statistics. *Sib. Zh. Ind. Mat.*, 4:2, 142–153 (in Russian).

[20] Zipf, G. K., 1936. *The Psycho-Biology of Language*. Routledge, London.

A scalar measure of interdependence between random vectors in problems for researching of multidimensional stochastic systems

A. Tyrsin[1,3], Ye. Chistova[2], A. Antonov[1]
[1] *Ural Federal B.N. Yeltsin University, Yekaterinburg, Russia*
[2] *Institute of Economics, Ural Branch of the Russian Academy of Sciences, Yekaterinburg, Russia*
[3] *Moscow Institute of Physics and Technology (National Research University), Moscow region, Dolgoprudny, Russia*
email: at2001@yandex.ru, elvitvas@ya.ru, antonovgr@mail.com

Abstract

A quantitative measure is described for assessing the tightness of interdependence between random vectors of different dimensions. It is expressed through the coefficients of determination of conditional regressions between components of random vectors. For the particular case of Gaussian random vectors, this value is expressed in terms of the determinants of each of the random vectors and the determinant of their union. It is shown that the coefficient introduced meets all the basic requirements for the degree of interdependence tightness between random vectors. This coefficient has advantages over the parameters obtained using canonical correlations analysis. It allows to determine the actual tightness of interdependence between random vectors. The measure introduced is fairly easy to interpret. The issues of its practical application in the economy are considered. The efficiency of social programs financing in the regions was evaluated.

Keywords: random vector, interdependence, stochastic system, differential entropy, index of determination, correlation matrix.

Introduction

When applying statistical analysis methods for the study of real objects, one often comes across their multidimensionality [1]. The main purpose of the multidimensional statistical analysis is to identify the nature and structure of the relationships between the components of the multidimensional feature under study [2]. One of the actual problems is the quantitative assessment of the tightness of the joint relationship (correlation) between multidimensional random variables. In many applications, there are several output variables. There arises a problem of assessing the relationship between the set of measured (input) variables $\mathbf{X}$ and the properties (output variables) $\mathbf{Y}$ of interest [1]. If the number of output variables is more than one, then the assessment of the tightness of the correlation between groups of input and output variables is carried out using canonical correlations analysis [3, 4]. This method is a generalization of the pair-wise linear correlation and allows finding most of the correlations between two groups of random variables. The dependence is estimated

using canonical variables calculated as linear combinations of the original features for each of the groups. These canonical values should correlate as much as possible among themselves, and their number is determined by the number of variables in the smaller set. This method has several significant drawbacks.

First, it is designed only for the case when all the studied features are connected linearly with each other, which actually implies the existence of a joint normal distribution for each of the random vectors. Second, the maximum value of the correlation coefficient between canonical variables is found, while it is required to estimate the tightness of the actual relationship, which can differ significantly from the maximum possible value. Together with the presence of a set of the determined correlation coefficients (their number is equal to the dimension of the output variables vector), this makes it very difficult to interpret the results, and it is too general and uninformative for practical purposes. Third, the representation of canonical variables only in the form of linear combinations of each of the variables groups limits the results of maximization. Fourth, it is required that the number of input variables should be not less than the number of output variables, which is also another limitation. It should be noted that the task of quantifying the tightness of the relationship between the components of a random vector has been solved. For Gaussian random vectors in [5], the coefficient of tightness of joint linear correlation is proposed

$$D_{e,m}(\mathbf{X}) = 1 - |\mathbf{R}_{\mathbf{X}}|^{1/m}, \tag{1}$$

where – $\mathbf{X} = (X_1,\ X_2,\ \ldots\ X_m)$ is a multidimensional random variable with a joint normal distribution and the correlation matrix $\mathbf{R}_{\mathbf{X}}$.

In [6], the interdependence tightness coefficient for linear and nonlinear case is introduced

$$d_{e,m}(\mathbf{X}) = 1 - \exp[-2I(\mathbf{X})/m], \tag{2}$$

where $I(\mathbf{X}) = H(\tilde{\mathbf{X}}) - H(\mathbf{X})$ – is the difference of differential entropies of multidimensional random variables $\mathbf{X}$ and $\tilde{\mathbf{X}} = (\tilde{X_1},\ \tilde{X_2},\ \ldots\ \tilde{X_m})$, whose components $\tilde{X_i}$ are mutually independent and have the same distributions as X_i.

The main disadvantage of formula (2) is the low accuracy of estimation of the coefficient $d_e(\mathbf{X})$, since to calculate the differential entropy $H(\mathbf{X})$, it is required to estimate, by bounded samples, one-dimensional and multidimensional densities of distributions of correlated random variables. For the purpose of practical implementation in [7], an equivalent formula was proposed instead of (2),

$$d_{e,m}(\mathbf{X}) = 1 - \Big[\prod_{k=2}^{m}(1 - R^2_{X_k/X_1X_2\ldots X_{k-1}})\Big]^{1/m}, \tag{3}$$

where $R^2_{X_k/X_1X_2\ldots X_{k-1}}$ – are determination indices of the corresponding regression dependencies, $k = 2, 3\ldots, m$. The issues of non-parametric estimation of the determination indices using multidimensional sample data are considered in [8]. The normalization (exponentiation of $1/m$) proposed in [5, 6] is not formally justified

even for the joint normal distribution. Indeed, for a Gaussian random vector $\mathbf{X}$ we have [8]

$$I(\mathbf{X}) = \sum_{i=1}^{m} H(X_i) + \frac{1}{2}\ln|\mathbf{R}_{\tilde{\mathbf{X}}}| - \sum_{i=1}^{m} H(X_i) - \frac{1}{2}\ln|\mathbf{R}_{\mathbf{X}}| = \frac{1}{2}\ln|\mathbf{R}_{\mathbf{X}}|,$$

therefore, the value $I(\mathbf{X})$ does not explicitly depend on the dimension m of the vector $\mathbf{X}$. Therefore, along with (1) and (3), the formulas are justified

$$D_e(\mathbf{X}) = 1 - |\mathbf{R}_{\mathbf{X}}|, \tag{4}$$

$$d_e(\mathbf{X}) = 1 - \prod_{k=2}^{m}(1 - R^2_{X_k/X_1X_2\ldots X_{k-1}}). \tag{5}$$

The purpose of the present research is to describe and test new quantitative feature of the interdependence tightness between two groups of variables, which does not have the disadvantages and limitations of the features inherent in canonical correlations analysis.

1 A measure of interdependence between random vectors of arbitrary dimensions

Next, we introduce the scalar measure of interdependence between continuous random vectors of arbitrary dimensions $\mathbf{X} = (X_1, \ldots X_m)$ and $\mathbf{Y} = (Y_1, \ldots Y_l)$.

We define vectors $\tilde{\mathbf{X}} = (\tilde{X}_1, \ldots \tilde{X}_m)$ and $\tilde{\mathbf{Y}} = (\tilde{Y}_1, \ldots \tilde{Y}_l)$ such that all components $\tilde{X}_i$ are mutually independent and have the same distributions as X_i, and all $\tilde{Y}_j$ – are mutually independent and have the same distributions as Y_j. We also define two random vectors of size $m+l$ as $\mathbf{Z} = \mathbf{X} \cup \mathbf{Y} = (X_1, \ldots X_m, Y_1, \ldots, Y_l)$ and $\tilde{\mathbf{Z}} = \tilde{\mathbf{X}} \cup \tilde{\mathbf{Y}} = (\tilde{X}_1, \ldots \tilde{X}_m, \tilde{Y}_1, \ldots, \tilde{Y}_l)$.

Let $I(\mathbf{X}, \mathbf{Y}) = I(\mathbf{Z}) - I(\mathbf{X}) - I(\mathbf{Y})$, where $I(\mathbf{Z}) = H(\tilde{\mathbf{Z}}) - H(\mathbf{Z})$, $I(\mathbf{X}) = H(\tilde{\mathbf{X}}) - H(\mathbf{X})$, $I(\mathbf{Y}) = H(\tilde{\mathbf{Y}}) - H(\mathbf{Y})$ – are differential entropy differences of random vectors. It is known [7] that $I(\mathbf{X}) = -\frac{1}{2}\sum_{k=2}^{m}\ln(1 - R^2_{X_k/X_1X_2\ldots X_{k-1}})$. From here we can get

$$I(\mathbf{X},\mathbf{Y}) = -\frac{1}{2}\ln(1 - R^2_{Y_1/X_1\ldots X_m}) - \frac{1}{2}\sum_{k=2}^{l}\ln\frac{(1 - R^2_{Y_k/X_1\ldots X_mY_1\ldots Y_{k-1}})}{1 - R^2_{Y_k/Y_1\ldots Y_{k-1}}}. \tag{6}$$

We introduce the interdependence tightness coefficient $d_e(\mathbf{X}, \mathbf{Y})$ between random vectors $\mathbf{X}$ and $\mathbf{Y}$ as $d_e(\mathbf{X}, \mathbf{Y}) = 1 - e^{-2/(\mathbf{X},\mathbf{Y})}$. Keeping in mind (6) we have

$$d_e(\mathbf{X}, \mathbf{Y}) = 1 - (1 - R^2_{Y_1/X_1\ldots X_m})\prod_{k=2}^{l}\frac{1 - R^2_{Y_k/X_1\ldots X_mY_1\ldots Y_{k-1}}}{1 - R^2_{Y_k/Y_1\ldots Y_{k-1}}}, \tag{7}$$

or

$$d_e(\mathbf{X}, \mathbf{Y}) = 1 - \frac{1 - d_e(\mathbf{X} \cup \mathbf{Y})}{(1 - d_e(\mathbf{X}))(1 - d_e(\mathbf{Y})}. \tag{8}$$

For the particular case, when the vectors $\mathbf{X}$ and $\mathbf{Y}$ have joint normal distributions instead of (8), we have

$$D_e(\mathbf{X}, \mathbf{Y}) = 1 - \frac{|\mathbf{R}_{\mathbf{X}\cup\mathbf{Y}}|}{|\mathbf{R}_{\mathbf{X}}||\mathbf{R}_{\mathbf{Y}}|}, \tag{9}$$

where $|\mathbf{R}_{\mathbf{X}}|, |\mathbf{R}_{\mathbf{Y}}|, |\mathbf{R}_{\mathbf{X}\cup\mathbf{Y}}|$ – are correlation matrix determinants.

The following theorem holds [9].

Theorem. The scalar measure $d_e(\mathbf{X}, \mathbf{Y})$ of the interdependence tightness between random vectors $\mathbf{X}$ and $\mathbf{Y}$ satisfies the following conditions:

1. $0 \le d_e(\mathbf{X}, \mathbf{Y}) \le 1$.

2. The case $d_e(\mathbf{X}, \mathbf{Y}) = 0$ corresponds to the independence between $\mathbf{X}$ and $\mathbf{Y}$, i.e. at least one of the components of vector $\mathbf{Y}$ is functionally (not randomly) associated with the components of vector $\mathbf{X}$.

3. Case $d_e(\mathbf{X}, \mathbf{Y}) = 1$ is assumed to have a functional relationship between $\mathbf{X}$ and $\mathbf{Y}$, i.e. at least one of the components of vector $\mathbf{Y}$ is functionally (not randomly) associated with the components of vector $\mathbf{X}$.

4. Feature $d_e(\mathbf{X}, \mathbf{Y})$ is a continuous function.

It should be noted that normalization can also be introduced for the coefficients $d_e(\mathbf{X}, \mathbf{Y})$ and $D_e(\mathbf{X}, \mathbf{Y})$. Since the dimensions m and l can change simultaneously, while maintaining a constant value $m + l$, we will normalize relative to the average dimension of the vectors equal to $(m + l)/2$. Therefore, we will assign the following expressions to the formulas (7)–(9):

$$d_{e,m,l}(\mathbf{X}, \mathbf{Y}) = 1 - \left((1 - R^2_{Y_1/X_1\ldots X_m}) \prod_{k=2}^{l} \frac{1-R^2_{Y_k/X_1\ldots X_m Y_1\ldots Y_{k-1}}}{1-R^2_{Y_k/Y_1\ldots Y_{k-1}}}\right)^{\frac{2}{m+l}},$$

$$d_{e,m,l}(\mathbf{X}, \mathbf{Y}) = 1 - \left(\frac{1-d_e(\mathbf{X}\cup\mathbf{Y})}{(1-d_e(\mathbf{X}))(1-d_e(\mathbf{Y}))}\right)^{\frac{2}{m+l}}, \; D_{e,m,l}(\mathbf{X}, \mathbf{Y}) = 1 - \left(\frac{|\mathbf{R}_{\mathbf{X}\cup\mathbf{Y}}|}{|\mathbf{R}_{\mathbf{X}}||\mathbf{R}_{\mathbf{Y}}|}\right)^{\frac{2}{m+l}}.$$

2 Testing the interdependence feature between random vectors by evaluating the effectiveness of social programs financing in the regions

Consider the use of interdependence feature between random vectors to assess the effectiveness of social programs financing in the regions of Russia from 2007 to 2016. The Human Development Index (HDI) was used as a general feature of regional development [10]. In order to take into account the regional specifics of the development, the Russian Federation constituents are divided into 3 groups based on the HDI value. To ensure statistical stability of the results when building a model, 17 constituent entities of the Russian Federation with the highest and the lowest HDI values, as well as with abnormal values or with missing statistical data on individual indicators were not taken into account. The model also includes features of social programs financing in the regions (expenditures of the Russian Federation constituent entity consolidated budget and of the regional state extra-budgetary fund on health care (X_1), education (X_2) and housing and utilities (X_3) – input features), and 9 resulting

features of regional development: life expectancy (Y_1); per capita income, rub. (Y_2); the annual cost of a fixed set of consumer goods and services, rub. (Y_3); population with income below the subsistence minimum,

A prerequisite for selecting the resulting features for the social programs financing of the regions, in addition to their widespread use as features of the quality of life (including in legal documents), was the existence of a statistical link between them and the HDI. A discriminant analysis was carried out [4] for annual time-series data of the resulting indicators Y_1, Y_2, ..., Y_9 for three groups ($IND = 1$, $IND = 2$, $IND = 3$) during the retrospective period. The results showed steady discrimination of the three groups.

The stability analysis of the regions in three groups and the results of evaluating the correlation tightness between the features of social programs financing in the region and those of the regional socioeconomic development allows us to make the following conclusions.

1. The tightness of the relationship between the features of social programs financing in the region grows with the increase of HDI, i.e. in developed regions, budget expenditures are more efficient.

2. The relationship tightness between the resulting indicators of socioeconomic development decreases with increasing HDI. This can be explained by a slowdown in the coordinated development of the regions. This conclusion is consistent with the instability of the regions in the same group, especially for entities II (with average HDI values) and entities III (with high HDI values) groups.

3. Thus, there are opposite trends in the relationship tightness between the features of social programs financing in the region and the resulting features of regional socioeconomic development with a change in the HDI.

Next, we determine the tightness of the correlation relationship between the financing of the health care system, education, housing and communal services, and the resulting features for the three groups of regions and the value averaged over the groups. The assessment results are presented in Figure 1.

In general, the results of the study showed a low efficiency of social programs financing. The socioeconomic development of the regions is rather determined not by the regional policy and the level of its financing, but by the level and characteristics of the region itself (availability of resources and favorable market conditions). Based on the estimates obtained, it can be concluded that the effectiveness of the social programs financing in the regions with different levels of socioeconomic development is clearly differentiated in times of economic instability. Anti-crisis financing aimed not at socioeconomic development, but at maintaining the current situation and mitigating negative consequences, is more effective for less developed regions of Russia. How it happened in 2009 and 2014. The opposite trend was observed in 2012 and 2016 regarding the prosperous regions of Russia. Therefore, as crisis phenomena are resolved in terms of the socioeconomic development of the underdeveloped region

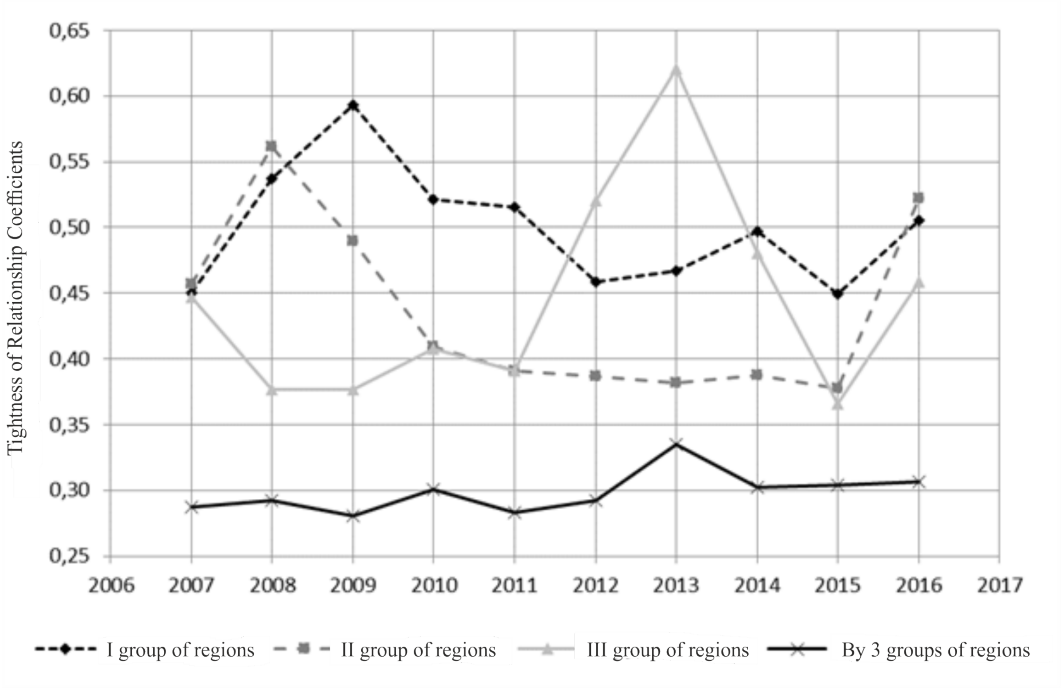


Figure 1: The results of the assessment of the relationship between the financing of the health care system, education and housing and communal services and the resulting features

(high unemployment rate, super high mortality rate, etc.), it becomes imperative to increase financing in social programs. However, this does not mean that only when the minimum criteria for socioeconomic development are achieved, it is necessary to finance long-term projects, they will simply be less effective. In stable periods of development (2007, 2016), the efficiency of the social programs financing in various regions is approximately at the same level. The presented calculations showed that in the regions with low HDI values (group I) the tightness of correlation between the sets of resulting and financial features is higher than the value obtained for more developed regions. With the increase in the HDI value of a region, the impact of budget financing on the quality of life of its population is reduced, which is explained by the decrease in the effectiveness of social programs financing in developed regions. As have already been noted by many researchers, the existing mechanism of stimulation of the Russian Federation entities that have reached the maximum level of socioeconomic development, without taking into account the additional parameters of funding efficiency, is not entirely fair.

Conclusions

1. A scalar measure of interdependence between continuous randomly distributed random vectors is introduced.

2. The partial results of the introduced measure for Gaussian random vectors are obtained.

3. The simplicity of the analysis obtained makes it possible to apply the proposed feature on sufficiently small samples.

4. The use of the interdependence feature between random vectors by evaluating the effectiveness of social programs financing in the regions is presented.

Acknowledgements

This paper was prepared with financial support from the Russian Science Foundation (Project No.18-010-000493).

References

[1] Esbensen K.H. (2002). *Multivariate Data Analysis – In Practice. – 5th Edition.* – CAMO Process AS, Oslo, Norway

[2] Prohorov Y.V. (1999). *Probability and Statistics: Encyclopedia.* – Big Russian Encyclopedia, Moscow, Russia, 910 p.

[3] Manly B.F.J., Navarro A.J.A. (2017). *Multivariate Statistical Methods. A Primer. – 4th ed. .* – CRC Press, 255 p.

[4] Soshinkova L.A., Tamashevich V.N., Webe G., Shefer M. (1999). *Multivariate statistical analysis in economics.* – UNITY-DANA, Moscow, 598 p.

[5] Pena D., Rodriguez J. (2003). Descriptive Measures of Multivariate Scatter and Linear Dependence. *Journal of Multivariate Analysis.* Vol. **85**, pp. 361-374.

[6] Pena D., Van der Linde A. (2007). Dimensionless Measures of Variability and Dependence for Multivariate Continuous Distributions. *Communications in Statistics: Theory and Methods.* Vol. **36**, pp. 1845-1854.

[7] Tyrsin A.N. (2014). A measure of the joint correlation of multidimensional random variables. *Factory laboratory. Diagnostics of materials.* Vol. **80**, pp. 76-80.

[8] Tyrsin A.N. (2016). *Entropy modeling of multidimensional stochastic systems.* Publishing house "Scientific Book", Voronezh, 156 p.

[9] Tyrsin A.N. (2018). The scalar measure of interdependence between random vectors. *Factory laboratory. Diagnostics of materials..* Vol. **84**, pp. 76-82.

[10] Chakravarty S.R. (2003). A Generalized Human Development Index . *Review of Development Economics.* Vol. **7**, pp. 99-114.

Data mining application features for scientific migration

Gavriil A. Agarkov, Anastasia E. Sudakova and Alexander A. Tarasyev
Ural Federal University, Ekaterinburg, Russia
e-mail: g.a.agarkov@urfu.ru, a-chusova@mail.ru, a.a.tarasyev@urfu.ru

Abstract

The study of scientific migration is of interest from the point of view of the analysis of economic behaviour of an individual, and the main problem for these studies is the lack of reliable data on migration of researchers. We have collected the relevant data using the author's methodology of data mining, which is implemented in the software based on big data technology. This methodology is based on the analysis of information presented in the abstract and citation databases of scientific literature: data on scientific migration are obtained by analysing changes in affiliation. The methodology was tested on the Scopus database. The practical value of the study lies in the possibility of using its results for the development of human resources of Russian science. The paper deals with the current trends and problems of scientific migration. According to statistics, the main factors of scientific migration were financial factors and improvement of working conditions in the centres of scientists attraction. It should be emphasized that scientific migration cannot be analysed only in terms of economic efficiency, but also takes into account the personal skills and preferences of scientists. The most effective phenomenon for economic growth is the brain sharing, when highly qualified scientists have the opportunity to travel abroad, improve their skills and at the same time participate in scientific projects of their main scientific organization.

Keywords: dynamic modelling, data mining, optimization, international scientific migration, scientific capital, academic mobility, human capital export, behavioural economics.

Introduction

Modern models of the development of the socio-economic system suggest a significant development of knowledge-intensive industries. At this stage of development, states and corporations invest huge amounts of money in the development of fundamental and applied science. Despite the high cost of developing the scientific infrastructure, the most scarce resource for increasing the effectiveness of science is human resources. Under the circumstances, the problem of scientific migration can not be considered without taking into account the global competition for scientists. It happens, both for well-known specialists with a strong background, and for young researchers starting their career. Competition for talent influences innovation policy initiatives around the world, stimulating the creation of new jobs and attracting foreign intellectual capital.

Attention to the study of scientific migration attracted the phenomenon, the so-called brain drain. Expanding the role of knowledge-intensive industries for economic development has increased the scientific interest in the problem of scientific migration. It is impossible to consider scientific migration in isolation from the issues of technology transfer. In this connection, the results obtained in the study of scientific mobility are interesting, indicating that organizations whose employees are involved in the "circulation of intelligence" between countries are the most productive in creating new technologies.

New theoretical approaches and concepts are being developed to evaluate the processes of scientific migration. For example, according to a number of experts, the current model of migration of scientists is characterized by a transition from a model of brain drain from one country and the growth of intellectual potential in another (brain gain) to the model of sharing intellectual potential (brain sharing). One of the conclusions about the migration to the migratory model is the thesis that international mobility benefits all parties, including countries that are net exporters of researchers.

A number of scientists believe that the priority motive of scientific migration is the search for an environment where they are most productive, while the environmental factors, individual indicators of scientific effectiveness, family factors are significant factors that influence the scientist's mobility. From our point of view, the policy of multiculturalism helps to attract cadres necessary for the country's development. Immigrants who remain in touch with their countries of origin can contribute to the reverse "brain drain". The methods used to study scientific migration are also of interest. For example, using methods of mathematical modelling, the authors identify factors affecting the plans of young scientists on immigration.

The application of theories of social inequality and decision-making in the field of education and migration within the life cycle framework is considered. The study of scientific migration is an extremely topical problem for Russia. In the 2010-ies among emigrants from Russia from 30 to 70 percent (depending on the country) were people with higher education. We recognize that the lack of information at the state level about emigrant scientists makes it difficult to understand the problem. Data are available only from sample surveys and studies that are of a local nature. The long-term leak of Russian scientists abroad is not just a problem, but at a certain stage it becomes a resource for the innovative development of the Russian economy.

1 Theoretical framework

1.1 Research problem

The study of migration of scientists is of scientific interest from the point of view of macroeconomic research due to the fact that the quality of human capital is the most important driver of the modern economy. However, migration is fraught with positive and negative impact.

One of the main trends of science is to increase the mobility of scientists. It is

assumed that this growth is one of the factors of scientific and technological progress. There is a common understanding that migration flows affect the scientific activities of a country, so both scientists and policy makers are interested in studying accurate models of this impact, the structure of flows, and predictors of the scale and direction of migration.

Although there are many studies on academic mobility, most of them are devoted to international migration [3], [14], [9], [10]. To learn how national science works, what are its strengths and weaknesses, it is really important to know about cross-border flows of scientists, but it is equally important to know how mobility is structured on a national scale.

The main difficulty in carrying out this kind of research is the lack of reliable and complete information about the migration of scientists, because official statistics give data on actually migrated individuals, and organizations do not form reports on the mobility of scientists, which would be presented in the public domain. The construction of internationally comparable mobility indicators for the scientific workforce is a persistent policy need.

1.2 Theoretical framework

Data on mobility largely comes from census data, labour force survey data recording (e.g. FCN), longitudinal panels (e.g. NCSES, Eurostat, ETC.), individual and organization studies, and case studies, none of which are considered sufficient to provide a comprehensive and up-to-date analysis of scientific migration for policy purposes [1], [6].

In most cases, studies are conducted on mobility in General for highly qualified personnel, without highlighting the features of sub-immigrant groups. Traditionally, the mobility of the scientist is studied using such methods as state statistics on the teaching staff and data on migration [2], analysis of summary and personal web pages [13], questionnaires and interviews with scientists themselves [4], [8], as well as government and administrative databases [7].

As an alternative, we propose to use bibliometric data analysis. Studying the distribution and trends in institutional affiliation, it is possible to track the trajectory of individual scientists (for example, through the author's profile, as on the Scopus database), as well as to analyze mobility at the level of academic groups, individual disciplines and countries [11]. This method allows to conduct a comparative study of the publishing activities of mobile and non-mobile authors Pearson and Cotgrave [12], and to study the impact of migration on the development of various disciplines Borjas and Doran [5]. Using the method of affiliation, it is also possible to study the mobility of groups of elite scientists, which are small, but nevertheless important for the development of science. These digital traces can record the movement of scientists between countries; concentrations of representatives of different disciplines in certain countries or organizations; and allow for the analysis of relative migration flows.

Given that publications are associated with date information, we can conduct diachronic network analyses to identify the trade of scholars not only between locations,

but among all locations over time. Publication data provide the added advantage that we can examine the impact of mobility, by measuring citations before, after, and during periods of transition. Furthermore, bibliometric data can be analyzed at least quarterly, which addresses the problem of the delays in obtaining statistics on R&D personnel that has been repeatedly noted in the literature.

The research of Moed and Halevi [11], who examined migration balances between a select group of developing and developed countries, using Scopus data. This work was methodologically useful in that it both discussed and examined the difficulties of author-name disambiguation for bibliometric data, including complexities of homonyms and synonyms in the database.

2 Data Mining Algorithm

2.1 Data Collection

The collection of data on academic mobility is carried out according to our program of searching and processing data, created on the basis of original algorithms proposed during the project. The initial data array for articles is presented in the form of a column-matrix:

$$A_{s1} = (a_{s1})_{r \cdot 1}, s = 1, ..., r, \tag{1}$$

where the column is the considered organization, s is the line with the article title, r is the total number of articles with affiliation to the analyzed organization.

These matrix arrays are collected automatically for each year. At the second stage, the A_{s1} matrix is expanded by adding new columns of characteristics to the analyzed data set: author's full name, affiliated country and organization, additional affiliations to other organizations and the number of citations for the considered article. As a result, 7 datasets on these indicators were obtained, reflecting Ural Federal University's scientific migration dynamics for the time period since 2011 up to 2017.

2.2 Data Processing

To reflect the information on the amount of articles, written by each author, affiliated to the chosen organization, we create a matrix B^k. Here we denote by i the parameter, which determines the author's ID. After that we launch the author search algorithm for the assembled data array, reflected by the matrix A, and obtain the elements b_{ij} for the matrix B^k by accumulating values in case the authors' identification numbers in the array coincide with the authors' target numbers. As a result of the transformations, we obtain the matrix B^k:

$$B^k = (b^k_{ij})_{m \cdot l}, \tag{2}$$

with the matrix lines representing all the authors of research papers, and the matrix columns $(i = 1, ..., m)$ - all identified academic organizations.

This matrix is compiled for each analyzed year, so each article has 3 main indexes that determine its position in the analyzed data array b_{ij}^k. We emphasize, that the resulting matrix B^k is a sparse matrix with elements reflecting information on the amount of articles b_{ij}^k of author i from organization j for selected year k.

In the compiled matrix B^k, the initial scientific organization is represented by the first column, so, if the affiliation to the main organization is identified, element b_{ij}^k is fixed in the first column, specifying the number of articles written by a sccientist, when he was involved into the first scientific organization's projects. When the algorithm detects a change in the author's profile for a new article, then the amount of new articles is assigned to the parameter b_{ij}^k, which is reflected in column j, where j is the column ordinal number indicating the organization whose article was published. To output data on brain drain based on the obtained data array, the elements of matrix B^k need to be checked for compliance with a number of conditions at the time interval k=[2011;2017]=[1;N].

2.3 Algorithm for data processing of scientometrics

In general, the solution of this problem can be represented by a conditional breakdown of the total time gap into two parts: $k_1 = [1; N/2]$ and $k_2 = [N/2{+}1; N]$ and derivation of the affiliation data to the considered Russian university in the first column $j = 1$. As a result, the task is reduced to the formation of a final matrix X reflecting the number of publications attributed to identified scientific organizations without reference to the time criterion:

$$X = (x_{ij})_{c \cdot d}. \tag{3}$$

On a time interval $k_1 = [1; N/2]$, the search and accumulation of an articles array attributed to authors i and scientific organizations j is carried out as follows: if for $\sum_{k=1}^{N/2} b_{ij}^k = x_{ij}$, $x_{ij} > 3$, then the author i is affiliated with the scientific organization j. Similarly, a search is performed on a time interval $k_2 = [N/2 + 1; N]$.

For each particular case, the following algorithm is checked:

- if the condition $\sum_{k=1}^{N/2} b_{ij}^k = x_{ij}$, $x_{ij} > 3$, j=1 is fulfilled on the time interval $k_1 = [1; N/2]$, the author is considered to be working in the main analysed organization (Ural Federal University);

- if the condition $\sum_{k=N/2+1}^{N} b_{ij}^k = x_{ij}$, $x_{ij} = 0$, j=1 is fulfilled on the time interval , the author is considered to be dismissed from the main analysed organization (Ural Federal University);

- if the condition $\sum_{k=N/2+1}^{N} b_{ij}^k = x_{ij}$, $x_{ij} = 0$, $j \neq 1$ is fulfilled on the time interval $k_2 = [N/2 + 1; N]$, the author is considered to be working in a foreign organization;

- if the condition $\sum_{k=N/2+1}^{N} b_{ij}^{k} = x_{ij}$, $x_{ij} = 0$, $j \neq 1$ is not fulfilled on the time interval $k_2 = [N/2+1; N]$, the author finished his scientific career.

As a result, the final matrix $X = (x_{ij})_{c \cdot d}$ reflects the number of articles written by the i-th author from the scientific organization j. To obtain an array of data on scientists who have moved to foreign universities, in the matrix $X = (x_{ij})_{c \cdot d}$ we strike out the first column, responsible for affiliation to the main analyzing organization. As a result, we obtain a sparse matrix $Y = (y_{ig})_{(c-1) \cdot d} = (b_{ij})_{c \cdot d} - (b_{i1})_{c \cdot 1}$, in which all authors who published their works from the UrFU and who have moved to foreign organizations are listed.

3 Results Discussion

The developed model and software were applied to analyze data on migration of scientists working at the Ural Federal University. During the collection and processing of data, a sample was obtained from 8986 scientists who published their work with affiliation with the Ural Federal University. A total of 367 scientists were identified, who experienced a constant change in the affiliation of research papers to foreign universities. After that, the data was further processed, their expert evaluation was carried out, data on scientists were questioned from the sample, and the change in affiliation was not related to emigration or another form of long-term cooperation with a foreign university or a research center.

Information was selectively verified according to information provided by scientists in social networks. The analysis of the revealed tendencies, in our opinion, is of interest for the research of the personnel potential of Russian science and education. It should be noted quite a wide geography of migration of Russian scientists, with the United States, Germany, Great Britain and France leading as the host countries (see Fig. 1).

Proceeding from the structure of scientific interests of scientists leaving Russia, it can be noted that the threat to the development of the Russian economy is not only the fact of their departure, but also the fact that among the departing the share of scientists engaged in the fields of knowledge is large, the results of studies in which either have a direct practical significance , or quickly find the introduction. In particular, such as engineering, computer technology, materials science. The demand for Russian scientists in the areas of science that provide rapid economic growth is evidenced by the fact that some scientists, while retaining their publications, continued their scientific and practical activities not in universities but in innovative companies. Including in such recognized world leaders of innovative economy, like Facebook, Google.

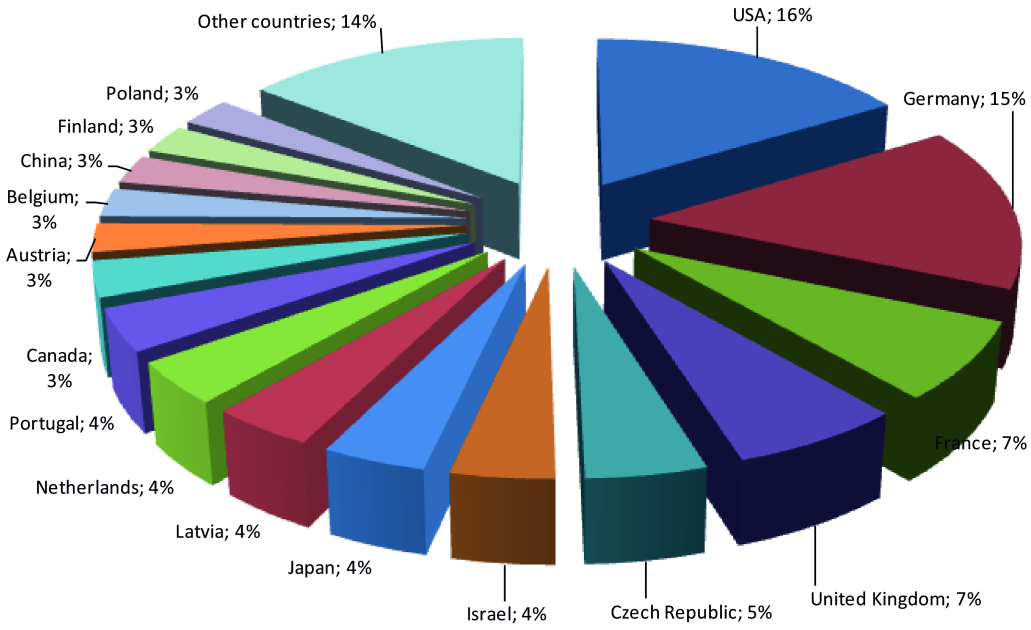


Figure 1: Countries hosting Russian scientists.

Conclusions

The study of the scientific migration of Russian scientists through the application of the methodology developed by the authors to analyze international reference databases of scientific publications confirms the significant impact of international migration of scientists on Russian science. Emigrate, both young scientists, and recognized by the scientific community of researchers. 18 percent of emigrated scientists have the Hirsch index in the range from 21 to 82. Migration of Russian scientists is directed mainly to countries of Western Europe and the United States (up to 72 percent). The main spheres of scientific interests of emigrating scientists are natural sciences and engineering (22 percent of physicists and astronomers, 16 percent of engineering specialists, 14 percent of materials scientists, 10 percent of mathematicians). Prospective directions for the development of research on scientific migration through the analysis of data in the reference databases of scientific publications is the expansion of the volumes of processed data sets presented in abstract research databases, for example, for all participants in the project to improve the competitiveness of Russian universities.

Acknowledgements

Grant of the President of the Russian Federation "Migration of Russian scientists in cross-border and national space: problems, efficiency, assessment" (MK-1737.2019.6)

References

[1] Åkerblom, M. (1996). Constructing internationally comparable indicators on the mobility of highly qualified workers. STI Review, 27.

[2] Arvizu D.E., Bowen R.M. (2014). National Science Board. 2014. Science and Engineering Indicators 2014. National Science Foundation, Arlington, VA.

[3] Bauder H. (2015). The International Mobility of Academics: A Labour Market Perspective. International Migration. Vol. 53(1), pp. 83–96.

[4] Boring P., Flanagan K., Gagliardi D., Kaloudis A., Karakasidou A. (2015). International mobility: findings from a survey of researchers in the EU. Sci Public Policy.

[5] Borjas G.J., Doran K.B. (2012). The collapse of the Soviet Union and the productivity of American mathematicians. Q J Econ. Vol. 127(3), pp. 1143–1203.

[6] Cacibano C., Bozeman, B. (2009). Curriculum vitae method in science policy and research evaluation: the state-of-the-art. Research Evaluation. Vol. 18(2), pp. 86– 94.

[7] De Filippo D., Casado E.S., Gymez I. (2009). Quantitative and qualitative approaches, to the study of mobility and scientific performance: a case study of a Spanish university. Res Eval. Vol. 18(3), pp. 191–200.

[8] Flanagan K. (2015). International mobility of scientists. The handbook of global science, technology, and innovation. Wiley, Chichester, pp. 364–381. doi:10.1002/9781118739044.ch17

[9] Franzoni C., Giuseppe S., Paula S. (2012). Foreign Born Scientists: Mobility Patterns for Sixteen Countries. Working Paper 18067. National Bureau of Economic Research.

[10] Gaule P. (2014). Who comes back and when? Return migration decisions of academic scientists. Economics Letters. Vol.124 (3), pp. 461-464.

[11] Moed H.F., Halevi G. (2014). A bibliometric approach to tracking international scientific migration. Scientometrics. Vol. 101(3), pp. 1–15.

[12] Pierson A.S., Cotgreave P. (2000). Citation figures suggest that the UK brain drain is a genuine problem. Nature. Vol. 407(6800), pp. 13–13.

[13] Sandstrm U. (2009). Combining curriculum vitae and bibliometric analysis: mobility, gender and research performance. Research Evaluation. Vol. 18(2), pp. 135-142.

[14] Stephan P., Giuseppi S., Chiara F. (2014). Migrant Scientists and International Networks: Diaspora and Home Country Advantage. As of 17 March 2017.

On segmentation approach for time series of arbitrary nature

A. Sherstobitova and T. Emelyanova
National Research Tomsk State University, Tomsk Russia
e-mail: annasherstobitova06@gmail.com, tv_em@mail.ru

Abstract

It is considered the problem of splitting the time series X_t of arbitrary nature into segments generated by one generation mechanism, as well as detecting the moment of change τ of one generation mechanism to another. Based on the parameters of ε-complexity, there is considered the time series segmentation methodology, which does not require any a priori knowledge of the mechanisms for their implementation.

Keywords: time series models, structural breaks, abnormal observation detection.

Introduction

There is often a situation in time series processing when time series are generated by various mechanisms, but a researcher has no a priori information about the mechanisms. In order to extract adequate information from a data array, it is necessary to segment the series into homogeneous (i.e. data sets, which are generated by the same mechanism) subarrays. Without this data preprocessing step, it is impossible to build mathematical models, estimate parameters, and so on.

When the data have been generated by a specific probabilistic mechanism, the segmentation problem is a problem of detecting structural breaks (moments of change in the mechanism of data generation) in random processes. Structural break detection is an essential and challenging task. To solve the challenge, we use an approach to the segmentation of time series of arbitrary nature proposed by B. S. Darkhovsky and A. Piryatinska [1]. The key idea of the approach is to adopt the ε-complexity which is an internal characteristic of the data and is sensitive to changes in the mechanisms of time series generation.

1 Segmentation Approach for Time Series

Let $X = \{x(t)\}_{t=1}^{N}$ be a time series with unknown moments of change in the generation mechanism $t_i, i = 2, \ldots, k$. he mechanisms of time series generation are unknown and can be stochastic, deterministic, or mixed. The segments of the series $[t_i, t_{i+1}], t_1 = 1, t_{k+1} = N$, which are generated by the same mechanism, are called homogeneous. B. S. Darkhovsky and A. Piryatinska formulated and proved the following theorem [1].

Theorem. For any function $x(\cdot)$ in a certain dense subset of the set of totally nontrivial functions satisfying the Hölder condition and defined by their n values on a

uniform grid and for any (sufficiently small) $\kappa > 0, \delta > 0$ and $n \geq n_0(x(\cdot))$, there exist a family of approximation methods F^*, numbers $0 < \alpha(n, x(\cdot)) < \beta(n, x(\cdot)) < 1$, $A(n, x(\cdot)), B(n, x(\cdot))$, $|B(\cdot)| \geq c(n, (x\cdot)) > 0$, functions $\rho(S), \zeta(S)$ and $N \subset Q = [\alpha(\cdot), \beta(\cdot)]$, $\mu(N) > \mu(Q) - \delta$ such that the following relations hold under approximation by the methods from $F \supseteq F^*$ for $S \in N$

$$\log \varepsilon = A + B + \rho(S) + \zeta(S), \sup_{S \in N} max(|\rho(S)|, |\zeta(X)|) \leq \kappa. \tag{1}$$

Thus, the relationship between the ε-complexity of a Hölder continuous function defined by an array of function values on a uniform grid and the approximation error ε is characterized by a pair of real numbers (A, B) denoting the complexity coefficients [1]. Taking a window of size n in accordance with the theorem, we calculate the complexity coefficients $R(j+1)$ for each segment of the series $x(t), t \in [jn+1, (j+1)n], j = 0, 1, \ldots, \left[\frac{N}{n}\right]$. We get a new diagnostic vector sequence $\{R(j)\}_j^{\left[\frac{N}{n}\right]}$ as a result.

The following hypothesis is the key idea of the proposed approach: for the i-th homogeneity segment $[t_i, t_{i+1}]$ of time series X_t for $t_i \leq t, t+n < t_{i+1}$, the complexity coefficients satisfy the relation

$$R(j) = R_i + \xi^i(j), \tag{2}$$

where R_i is the mean of the sequence $R(j)$ on the interval $[t_i, t_{i+1}]$, and $\xi^i(j)$ is a sequence of random variables with zero expectation. In other words, the hypothesis means that the mean values of the ε-complexity coefficients of the time series remain constant at homogeneity intervals, and a change in the mechanism of series generation at moments $t_i, i = 1, 2, \ldots, k$, leads to the change of these mean values.

Thus, if the hypothesis is true, the problem of time series segmentation is reduced to the problem of detecting structural breaks in the mean value of the diagnostic vector sequence $R(j)$. To solve the aforementioned problem, B. S. Darkhovsky and A. Piryatinska proposed to use the following family of statistics

$$Y(s, \delta) = (\frac{(N_1 - s)s}{N_1^2})^{\delta}(\frac{1}{s}\sum_{k=1}^{s} z(k) - \frac{1}{N_1 - s}\sum_{k=s+1}^{N_1} z(k)), \tag{3}$$

where $0 \leq \delta \leq 1, 1 \leq s \leq N_1 - 1, N_1 = \left[\frac{N}{n}\right], Z = \{z(k)\}_{k=1}^{N_1}$ is the implementation of the diagnostic sequence components R [2].

We illustrate the theorem and approach to the segmentation of time series proposed by B. S. Darkhovsky and A. Piryatinska. Piecewise polynomial functions up to the 10^{th} order inclusive as the family F of approximation methods were used. For each experiment, a time series was generated and then divided into "windows " of size n. Each "window" was processed by the method of least squares, and the coefficients of dependence (1) were determined by values of $\log S$ and $\log S$. Further, for each experiment, the moment of change in the generation mechanism of the series (structural breaks) was detected. The structural breaks were determined with the use of an algorithm based on statistics (3).

2 Results of Monte-Carlo Simulation

Let us consider the results of Monte-Carlo simulations. Let x_t be the process with three times of the change in the generation mechanism. Before the first structural break x_t is an autoregressive moving average process ($ARMA(2,1)$) with parameters $\phi = (0.003; 0.1), \theta = 0.002, \varepsilon_t$ are the independent identically distributed random variables, $E\varepsilon_t = 0, E\varepsilon_t^2 = 1$. After the first structural break, the process is described by a deterministic equation. After the second structural break, the process is described by an autoregressive moving average ($ARMA(2,1)$) model with the same parameters. After the third structural break, the time series is represented by independent identically distributed random variables $\xi_t \sim U[0;1]$.

$$x_t = \begin{cases} \varepsilon_t + sum_{i=1,2}\phi_i x_{t-i} + \theta\varepsilon_{t-1}, & t \le k_1 \\ x_{t-1}^2 - 2, & k_1 < t \le k_2 \\ \varepsilon_t + \sum_{i=1,2}\phi_i x_{t-i} + \theta\varepsilon_{t-1}, & k_2 < t \le k_3 \\ \xi_t, & t > k - 3. \end{cases} \tag{4}$$

Let us consider the illustration of the theorem (Figure 1). Here the x-axis is the logarithm of the number of dropped points in percentage form, and the y-axis is the logarithm of the error in the approximation of the resulting function by the method from the family F.

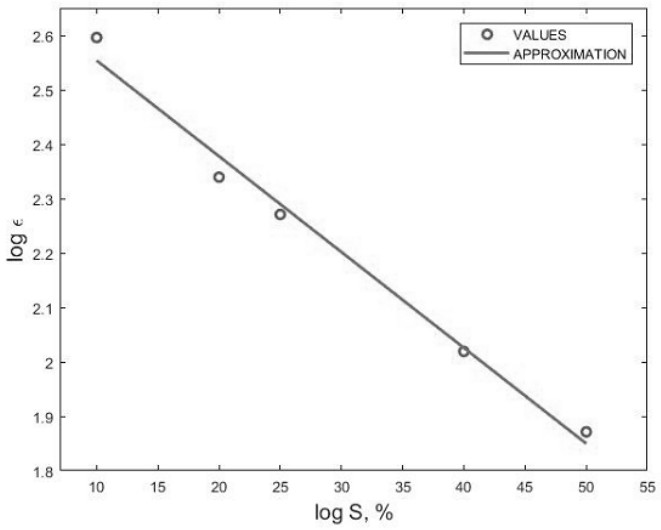


Figure 1: Illustration of the Theorem

Note that despite a relatively poor set of approximation methods F, the dependence of the form (1) is realized with good accuracy. Let us consider the results of Monte-Carlo simulations for the process (4) (Figure 2). The time series is generated by four different generation mechanisms and has three structural breaks: $k_1 = 310, k_2 = 620, k_3 = 1201$.

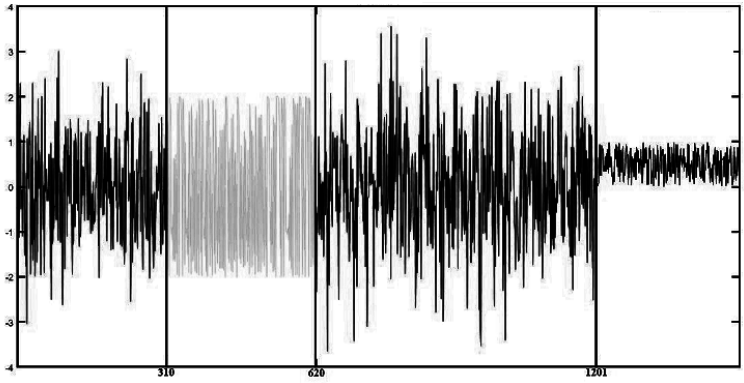


Figure 2: Illustration of the Model Realization (4)

The data sample size is $N = 1500$. The time series is divided into "windows " of the same size $n = 100$. A structural break is detected when the difference in the mean diagnostic sequence is significant. Results of Monte-Carlo simulations allow to detect structural breaks in 3, 6, and 12 windows, which corresponds to the intervals of 300-400, 600-700 and 1200-1300 in the time series x_t. Structural breaks are detected correctly.

In practice, it is often necessary to know the number of observations when the generation mechanism is changed, or abnormal data behavior is recorded, which is not typical for the time series observed. In this regard, a modification of the segmentation method of B. S. Darkhovsky and A. Piryatinska is proposed, which allows a specific number of observations with a fixed anomaly to be obtained.

Consider the model ARCH(2)

$$X_t = \varepsilon_t \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i X_{t-i}^2}, \tag{5}$$

where $\varepsilon\ N(0,1), \alpha = (0.5, 0.3, 0.6)$. The sample size for the experiment is $N = 1000$, and the window size $n = 100$.

The resulting series has a non-uniform behavior, so the anomalies are selected jumps being much higher than the average value of the series (Figure 3). We apply the segmentation method of B. S. Darkhovsky and A. Piryatinska to the time series, and, as a result, obtain 2, 3, 5, and 6 windows with possible anomalies. Next, we consider each identified window in the intersection with the neighbors as follows

$$X_1 = \{x_i \in X : i = \overline{nt - 50; n(t+1) + 50}\}, \tag{6}$$

where t is the identified window number. Thus, we consider four obtained time series to detect an abnormal observation. It is necessary to consider each observation in the new generated series separately and to check the following three conditions

$$X_{1_i} > \overline{X}_{1_{i-1}} + l; X_{1_i} > X_{1_{i-1}} + l; X_{1_i} > X_{1_{i+1}} + l, \tag{7}$$

where l is the controlling coefficient.

Thus, if the observation in the window satisfies three conditions in (7), then the observation is considered as abnormal.

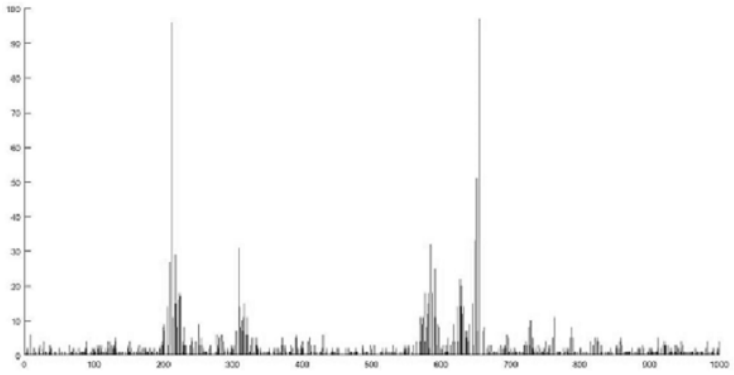


Figure 3: Detection of Abnormal Observations

Conclusions

Thus, the segmentation algorithm of B. S. Darkhovsky and A. Piryatinska allows not only time series to be segmented into subarrays generated by various generation mechanisms, but also abnormal observations to be detected in the time series by performing a simple modification of the algorithm. The modification proposed was tested not only on simulation data but also on real data. The modified segmentation method showed high accuracy in working with real data, which is an undeniable advantage.

References

[1] Darkhovsky B. S., Piryatinska A. (2014). New approach to the segmentation problem for time series of arbitrary nature. *Proceedings of the Steklov Institute of Mathematics*. Vol. **287**, pp. 54-67.

[2] Brodsky B. E., Darkhovsky B. S. (2000). *Non-parametric statistical diagnosis: Problems and methods*.Dordrecht: Kluwer.

Computer analysis and interpretation of geophysical data

Dmitriy S. Rusin, Natalia V. Molokova, Andrew G. Feldman, Natalia V. Nikolaeva
Siberian Federal University, Krasnoyarsk, Russia
e-mail: rusin199812@mail.ru, Nat_molokova@mail.ru, anrfeldman@yandex.ru, nikolnat@mail.ru

Abstract

In this paper, we consider the development of a software module for computer analysis and interpretation of the geophysical data of the studied area of the Barobinskoe field, used in prospecting and subsequent exploration of mineral deposits.

Keywords: Data analysis, effective resistance, envelope graphics, pickets .

Introduction

The use of modern methods of interpreting geological data is increasing every year. This is influenced by two factors: economic and temporal. Computer analysis and interpretation of geophysical data of the studied areas is of practical importance for the search and subsequent exploration of mineral deposits. Our goal is to reduce time input and eliminate manual processing errors thus advancing efficiency of geological prospecting. To achieve the stated goal, the following tasks were set: analyze the subject area, computerize processing field data of the long-wire method [1], and take into account ergonomic and psychological factors.

1 Domain Analysis

The analysis of the subject area showed that in the long-wire method for specifying an electromagnetic field a wire with a length of 1 to 10 km or more is laid. A magnetic and electric component of this field is recorded by a ferrite antenna (magnetic version) or a MN line (electric version) [2]. When passing an alternating current through a wire, the electromagnetic fields also create A and B groundings, which add up to the wire field and complicate the general picture. These complicating fields are local, most intense in places of grounding and quickly fall off to the periphery. To eliminate their influence, the working area of the survey is assumed to be equal to 0.8 of the AB line length. The "long wire" term refers to the length of the main, in which the grounding electrodes do not affect the field within the area under study. The boundaries of the section that meets the long-wire method conditions with an error of $\leq 5\%$ are shown in Fig. 1.

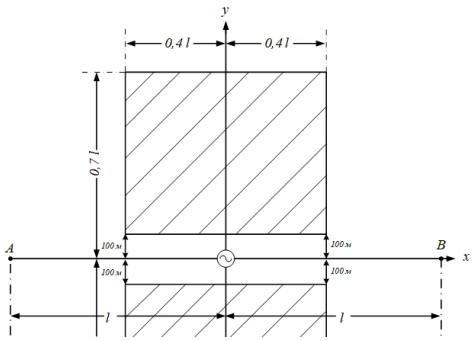


Figure 1: The range of the H_y and H_z values that meet the long-wire method conditions

2 Data sampling

A field survey of the magnetic field is carried out on either side of a long wire grounded at the ends, through which an alternating current of a fixed frequency is allowed. To achieve a great depth of study, a low frequency range is applied. The receiving device records the parameters in the long-wire field – horizontal H_y and vertical H_z, which make up the magnetic field, and calculates their ratio H_y/H_z, using which the effective resistance r_{eff} is defined with a pattern taking into account the operating frequency f and the spread r. r_{eff} of the electromagnetic method equipment is defined in the measuring device. These parameters as well as the picket values and power line positions are the source information, which is recorded from the equipment and stored in an Office MS Excel document. The source data contains information on more than 20 profiles. A fragment of the source numerical geophysical data is presented in Figure 2.

Date	Profile No	Picket	X	Y	P2	Hz	Hr	Hz/Hr	up to AB	ref
		7.1600	692839	5794917		69.92	226.88	0.308	0.4000	1316,0
		7.1640	692879	5794919		77.81	231.09	0.337	0.3600	864,3
		7.1680	692919	5794922		87.34	259.69	0.336	0.3200	682,9
		7.1720	692959	5794923		125.78	291.64	0.431	0.2800	317,1
		7.1760	692999	5794923		147.50	363.12	0.406	0.2400	253,8
		7.1800	693040	5794924		160.62	463.12	0.347	0.2000	253,1
		7.1840	693081	5794927		203.05	623.75	0.326	0.1600	183,1
		7.1880	693121	5794926		215.62	885.00	0.244	0.1200	215,3
		7.1920	693161	5794930						
		7.1960	693201	5794931						
	A1B1	7.2000	693242	5794933	supply line position					
		7.2040	693282	5794936						
		7.2080	693321	5794936						
06.04.2017		7.2120	693361	5794938		219.84	1032.50	0.213	0.1200	284,3
		7.2160	693400	5794938		205.23	823.75	0.249	0.1600	347,1
		7.2200	693441	5794940		200.31	692.50	0.289	0.2000	382,6
		7.2240	693481	5794943		190.47	605.00	0.315	0.2400	473,8

Figure 2: The source geophysical data

The source information is also stored graphically. For example, the graphs of the effective resistance values for profile No. 7 are shown in Fig. 3. Further over these schedules processing will be done. The graphs are presented for ease of application, and they fully correspond to the values of numerical data (Fig. 2).

The output is a geoelectric section of the study area. As a result of graphic interpretation, a section of contours of effective rock resistance is constructed, where

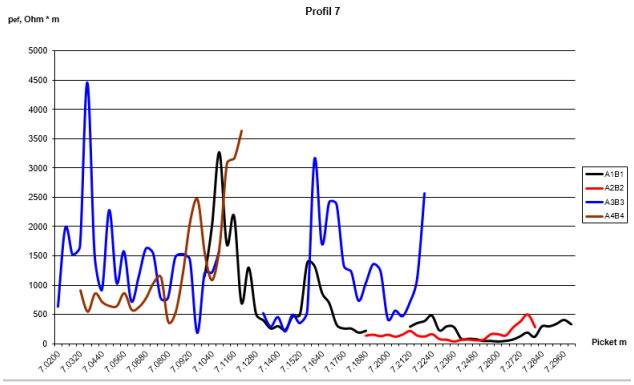


Figure 3: The graphs of the effective resistance values for profile No. 7

low and high resistance zones are distinguished with their substance prediction. To obtain the desired result, we have decided to develop a software module that would allow us to read geophysical data from the source file and automatically process it, but first we need prepare the input data. To do this we have to transfer the data shown in Figure 4 to an empty excel sheet.

Date	Profile No	Picket	X	Y	P2	Hz	Hr	Hz/Hr	up to AB	ref
		7.1600	692839	5794917		69.92	226.88	0.308	0.4000	1316,0
		7.1640	692879	5794919		77.81	231.09	0.337	0.3600	864,3
		7.1680	692919	5794922		87.34	259.69	0.336	0.3200	682,9
		7.1720	692959	5794923		125.78	291.64	0.431	0.2800	317,1
		7.1760	692999	5794923		147.50	363.12	0.406	0.2400	253,8
		7.1800	693040	5794924		160.62	463.12	0.347	0.2000	253,1
		7.1840	693081	5794927		203.05	623.75	0.326	0.1600	183,1
		7.1880	693121	5794926		215.62	885.00	0.244	0.1200	215,3
		7.1920	693161	5794930						
		7.1960	693201	5794931						
	A1B1	7.2000	693242	5794933	supply line position					
		7.2040	693282	5794936						
		7.2080	693321	5794936						
06.04.2017		7.2120	693361	5794938		219.84	1032.50	0.213	0.1200	284,3
		7.2160	693400	5794938		205.23	823.75	0.249	0.1600	347,1
		7.2200	693441	5794940		200.31	692.50	0.289	0.2000	382,6
		7.2240	693481	5794943		190.47	605.00	0.315	0.2400	473,8

Figure 4: Data selection for processing

The picket parameters are calculated by processing each filled cell from Excel, and then added to the two-dimensional array. The first value indicates the calculated picket position, the second – the effective resistance. Figure 5 shows an example of input parameters. Cell *A1* shows the power line picket, and then the wire pickets themselves are located in Column *A*. Each cell contains the profile number and, for each wire position, indicates the picket position on the profile in 40-meter increments. Column *B* locates the wire r_{eff}. As we use several wires, data on each wire is placed on different sheets, for convenience in operation.

3 Plotting envelope graphics

The next step was to build an envelope of the effective resistance graph. In a conducting anomalous object under the action of an alternating magnetic field of the wire

7.2000	
7.1000	1147,7
7.1040	2021,4
7.1080	3263,3
7.1120	1698,5
7.1160	2176,3
7.1200	695,7
7.1240	1295,7
7.1280	503,6
7.1320	393,4
7.1360	253,9
7.1400	294,1
7.1440	231,7
7.1480	460,1
7.1520	494,4
7.1560	1384,7
7.1600	1316,0

Figure 5: The input parameters

eddy currents are induced, which in turn generate a secondary magnetic field. The total magnetic field from conducting anomalous objects creates an anomaly of the measured field values at the observation sites, which is displayed in the r_{eff} graphs. Figure 7 shows the program output for constructing a graph of the effective resistance r_{eff} and the envelope. The envelope in this case is the arithmetic middling of all the values of each wire in the range from 2 to i-2, where i is the number of all the wire points in the input parameters.

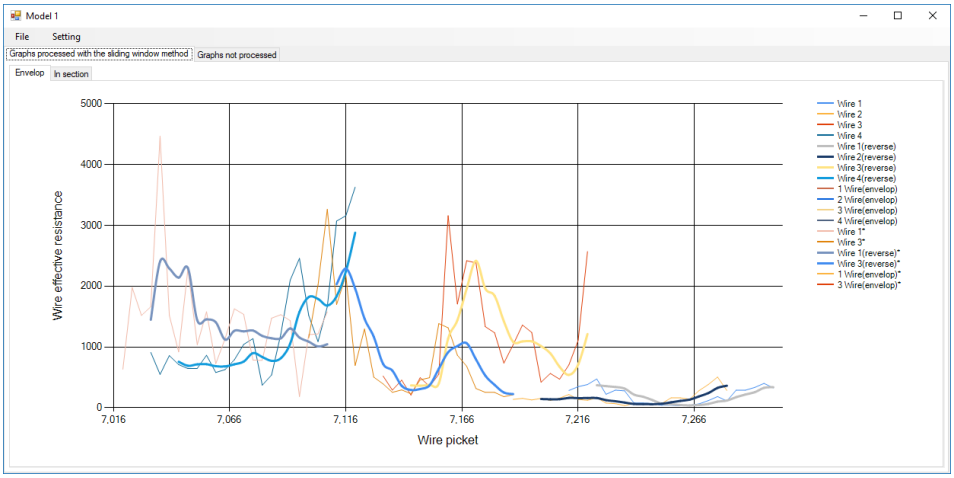


Figure 6: The program output

In this paper the calculated graph is presented for visual comprehension and is an intermediate action for the next work stages.

4 Plotting by method sliding window

The next step of our work was to construct a graph with the sliding window method in section conducted to eliminate the influence of near-surface heterogeneity. In the graph r_{eff} from the wire to the periphery, all the objects of shallow bedding

(tens of meters) are marked by local anomalies, which corresponds to the profiling function that solves the problem of geographic mapping of the studied area. If we remove these shallow anomalies (high-frequency background), the graph will have wide anomalies of deep-laid objects (hundreds of meters), which corresponds to the function of sounding.

The r_{eff} value at each observation site is formed depending on the magnitude and magnetic field direction H at each site of the geological environment under study, and the H, in turn, is an integral support of the eddy current caused in each conducting fragment of the environment by the primary alternating magnetic wire field and superimposed conduction and competence currents. Each elementary area (electromagnetic cell) of the lower half-space contributes to the total anomalous effect at the observation site, inversely proportional to its remoteness from current sources (wire and earth conductors) and remoteness from the observation site from the cell.

Since the geological structure of the section is individual at each wire side, then, following the semi-quantitative interpretation (containing considerable possible construction errors), the r_{eff} values calculated at the profile points are transferred to the intersection with the raypath drawn from K (wire) at an angle of 35° to the horizon (at provided $h = r/3$). In fact, this angle depends in ratio on the conductivity response and may vary within a broad range.

The values of the calculated picket parameters are plotted on the z axis, and effective resistance is plotted on the y axis.

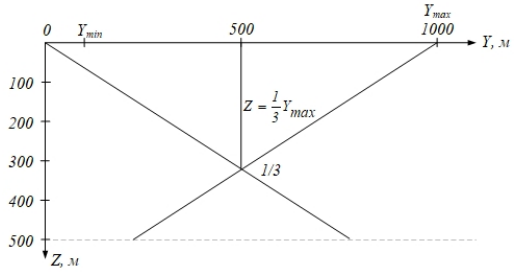


Figure 7: The calculated geometrical situation with the long-wire method (vertically). 0 – the wire position perpendicular to the drawing plane; Y_{min} and Y_{max} – minimum and maximum spread at the observation profile

To obtain the Z and Y values, the following formulas were applied:

$$Z = (Y_{min} + Y_i)/2. \tag{1}$$

$$Y = |Y_{min} - Y_i| * \frac{1}{3}. \tag{2}$$

For the value indicated in the point itself, we applied the formula (3):

$$X_i = Y_i. \tag{3}$$

The sliding window method consists in replacing the actual values of the series terms by the arithmetic mean of several terms closest to it. The set of averaged values

creates a sliding window. The term, which value is replaced with the mean with respect to the window, occupies the middle position in the window. The graph processed by the sliding window method in section can be seen in Figure 10.

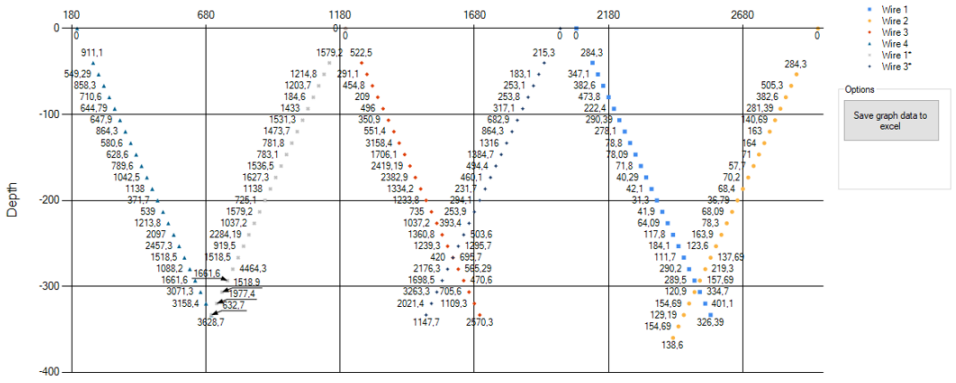


Figure 8: The graph of the values processed with the sliding window method in section

The picket parameters are calculated by processing each filled cell from Excel and added to the two-dimensional array. The first value indicates the calculated picket position, the second – the effective resistance.

The obtained values of the parameters with the sliding window method were stored in the Excel file for contouring. An example of the stored data (the sliding window method in section) is shown in Figure 11.

500	-333,333	1147,7		1000	1147,7
480	-320	2021,4		1040	2021,4
460	-306,667	3263,3		1080	3263,3
440	-293,333	1698,5		1120	1698,5
420	-280	2176,3		1160	2176,3
400	-266,667	695,7		1200	695,7
380	-253,333	1295,7		1240	1295,7
360	-240	503,6		1280	503,6
340	-226,667	393,4		1320	393,4
320	-213,333	253,9		1360	253,9
300	-200	294,1		1400	294,1
280	-186,667	231,7		1440	231,7
260	-173,333	460,1		1480	460,1
240	-160	494,4		1520	494,4
220	-146,667	1384,7		1560	1384,7
200	-133,333	1316		1600	1316
180	-120	864,3		1640	864,3
160	-106,667	682,9		1680	682,9
140	-93,3333	317,1		1720	317,1
120	-80	253,8		1760	253,8
100	-66,6667	253,1		1800	253,1
80	-53,3333	183,1		1840	183,1
60	-40	215,3		1880	215,3

Figure 9: An example of the stored data of the sliding window method in section

The columns *E, F* show the input values, the columns *A, B, C* show the processed graph data. Further, the data obtained by the sliding window method are processed in the Surfer software environment. As a result of processing, the surface effective resistance values are placed in depth, and a contour section is constructed, giving an approximate idea of the geoelectric section. The processing result is shown in Fig.

12. After that, the resistance contour concentrations are analyzed and evaluated for the desired minerals in them.

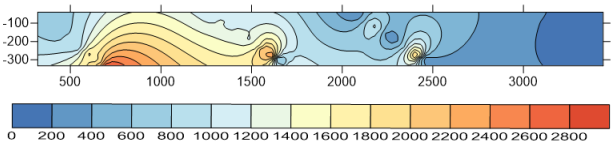


Figure 10: The resistance contours constructed in Surfer

Conclusion

High-resistance (1 and 2) and low-resistance (3) zones with their substance prediction are distinguished in this pseudosection. At the surface in the 800 - 1600 m spacing of the profile, fragments of quartz with gold were found, according to which it is assumed that we may find a series of gold-bearing quartz veins and stock work at the depth. In the 2000–3500 m spacing there is a low, marshy land, indicating the presence of salt marsh, which led to an extensive low-resistivity Zone 3, of little promise for geological prospecting and exploration. All three zones are recommended for well drilling, and when confirming the geophysical prediction Zones 1 and 2 – for prospecting. The material obtained is tentative data for the subsequent directed geological prospecting and, in general, significantly reduces its volume and cost.

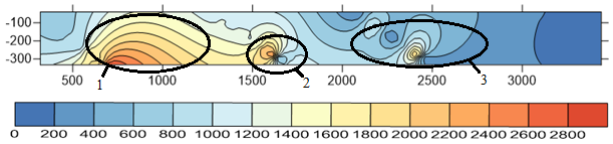


Figure 11: Substance evaluation

Figure 13 shows the fragments of quartz in selected Zones 1 and 2, as well as the predominance of wet clay (Zone 3) and the entire right-hand side.

Was developed the software module interface that allows the pickets to be calculated. When developing the interface, ergonomic factors were taken into account [3]. The interface allows easily loading data and processing it.

The graphs are navigated through the tabs located in the main program window, shown in the figure below. We can also use the manual setting buttons on the right sidebar for more precise setting. Each of these buttons will increase / decrease the height / width of the graph by the number indicated on the button being pressed.

In this paper psychological factors were taken into account. For this purpose, a user manual [4] was developed, which contains information about the program,

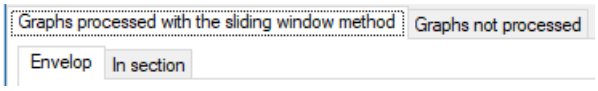


Figure 12: The main program window tabs

a detailed description of all the program features that are necessary for program authoring.

The developed software module will not only computerize processing geophysical data taken with the long-wire method, but also ensure safe storage and easy access to the information received. It also significantly reduces the time input for data processing, thereby reducing the cost of all the work done.

References

[1] V.I. Igolkin, A.V. Samkov, R. E. Toyb, M.F. Khokhlov. The equipment for the alternating harmonic current induction methods for mining and engineering geology (Apparatura induktivnikh metodov peremennogo garmonicheskogo toka dlya rudnoy i ingenernoy geologii) // Materials of All-Russian conference "Geophysical methods for exploration". – Tomsk: Publishing office of the TPU, 2016. – P. 219-222.

[2] Alternating-current prospecting methods and equipment (Metody i apparatura elektrorazvedki na peremennom toke): scientific publication / V.I. Igolkin, G.Y. Shaydurov, O.A. Tronin, M.F. Khokhlov. – Krasnoyarsk: SibFU, 2016. – 272 p.

[3] D.S. Rusin. The software module for processing geophysical data taken with the long-wire method (Programmniy modul dlya obrabotki geofizicheskikh dannikh, snyatikh metodom beskonechnogo dlinnogo kabelya) / D.S. Rusin / scientific supervisor N.V. Molokova. // Education and science in Russia and abroad, 2019, No 4, Vol. 52. – 518 p., P. 65-68.

[4] D.S. Rusin, N.V. Molokova. Control of the software module for processing geophysical data taken with the long-wire method (Upravlenie programmnoy modeliu dlya obrabotki geofizicheskikh dannikh, snyatikh metodom beskonechnogo dlinnogo kabelya) // E-Scio [electronic resource] – Electronic periodical "E-Scio.ru" – El No FS77-66730 — available at http://e-scio.ru/wp-content/uploads/2019/03/Русин-Д.-С.-Молокова-Н.-В..pdf

Statistical approach to detection of periodic signals under the background noise using the chaotic oscillator Murali-Lakshmanan-Chua

Tatiana V. Patrusheva and Egor M. Patrushev
Altai State Technical University, Barnaul, Russia
e-mail: attractor13@gmail.com, attractor@list.ru

Abstract

The article presents a method for detecting periodic signals under the background noise using the Murali-Lakshmanan-Chua non-autonomous chaotic oscillator. The authors propose to use the calculation of the number of chaotic bursts during the detection time to recognize the steady state of motion in the chaotic oscillator. The decision to detect a signal is made by comparing the received number with a threshold established by the Neumann-Pearson criterion. The authors hypothesized that the experimental data for the number of chaotic bursts has a negative binomial distribution and verified using the χ^2 criterion. A numerical model is presented that implements a system of differential equations for a signal detector based on chaotic oscillator Murali-Lakshmanan-Chua. The proposed model allowed us to perform statistical studies of the chaotic oscillator under the narrowband noises.

Keywords: bifurcations, signal detection, Murali-Lakshmanan-Chua oscillator, negative binomial distribution.

Introduction

Currently, a method of detecting deterministic signals under the background noise using bifurcations in a non-autonomous chaotic oscillator is known [3]. The basic principle of detection is as follows. Chaotic oscillator parameters are set in such a way that there is a chaotic state with intermittent behavior. Such a motion is possible near the boundary with tangential bifurcation. The received signal is fed into the chaotic oscillator and under the action of the detected signal a bifurcation occurs, and the chaotic state is replaced by a periodic steady state. One of the most developed approaches is to use the Duffing-Holmes oscillator in the numerical model for detecting periodic signals, and the recognition of the steady state is performed by calculating the largest Lyapunov exponent [4].

In creating miniature sensors [6], it becomes necessary to have a simple detector without using numerical models. As a chaotic oscillator for the detector, it is proposed to use the Murali-Lakshmanan-Chua oscillator [5], which is simpler than the Duffing-Holmes oscillator [7] realization in the form of an electrical circuit. To recognize the motion state, it is proposed to calculate the number of chaotic bursts during the detection time.

1 Basic principle

The construction of the detector is carried out on the basis of the Murali-Lakshmanan-Chua oscillator, which is described by the following system of equations in dimensionless form:

$$\begin{cases} \dot{x} = y - m(x) \\ \dot{y} = -\beta y - \beta x + A_0 \sin \omega\tau + \eta(\tau) \end{cases}, \tag{1}$$

where x and y are dynamic variables, β is a bifurcation parameter of the system, A_0 is an amplitude of the sinusoidal source, ω is a frequency of the sinusoidal source, τ is time, $\eta(\tau)$ is received signal, $m(x)$ is a function describing a non-linear element also known as Chua's diode:

$$m(x) = bx + \frac{1}{2}(a-b)(|x+1| - |x-1|), \tag{2}$$

where $a = -1.02$ and $b = -0.55$ are constant coefficients. The tangential bifurcation in this system is observed at $\beta = 0.9$, $\omega = 0.4$, $A_0 = A_{cr} \approx 0.0825$, where A_{cr} - critical value of amplitude. Near the tangential bifurcation an intermittent motion is observed in the system, consisting of long laminar phases separated by short chaotic bursts (Fig. 1). The appearance of chaotic bursts is random in nature. Thus, the

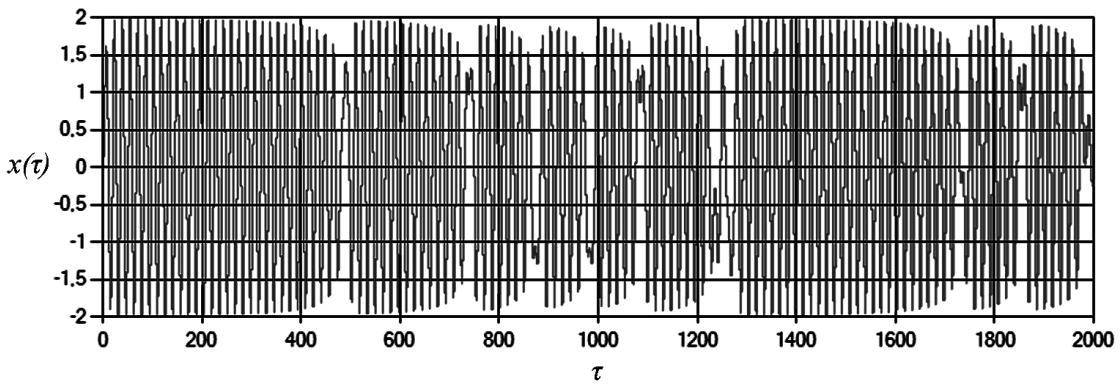


Figure 1: The time dependence for dynamic variable $x(\tau)$ with $\beta = 0.9$, $\omega = 0.4$, $A = 0.0824$

amplitude of the sinusoidal source A_0 must be set slightly below the bifurcation boundary A_{cr}, and a bifurcation must occur upon receipt of a detectable signal. Above the bifurcation boundary, the chaotic bursts cease. For the signal, supplied to detect, we introduce the following notation:

$$\eta(\tau) = d \sin \omega\tau + \sigma_{in} n(\tau), \tag{3}$$

where d is the amplitude of the detected signal $\sin \omega\tau$, σ_{in} is the effective value of the noise, $n(\tau)$ is a narrow-band Gaussian random process with zero mean and variance equal to one. Denote as A the sum of the amplitudes of the sinusoidal source and the detected signals. In order for a bifurcation to occur, inequality must be fulfilled:

$$A = A_0 + d > A_{cr}.$$

It is known that in chaotic oscillator, even with $A > A_{cr}$, chaotic bursts are possible under the action of noise [2]. To build a periodic signal detector, it is necessary to compare the number of chaotic burst N during the detection time τ_d with a certain threshold h', but setting this threshold requires knowledge of the distribution law of the random variable N.

2 Detection model

The signal detector under the background noise can be built according to the scheme shown in Fig. 2. The following notation is used: CO is a chaotic oscillator, PD is a binary phase detector, C is a counter, TD is a threshold device. The detector works as follows. The received signal $\eta(\tau)$ is fed to the input of a non-autonomous chaotic oscillator CO. Also, for a chaotic oscillator, the source of a sinusoidal signal $A_0 \sin \omega\tau$ is necessary. The output signal of the chaotic oscillator $x(\tau)$ is fed to the input of the binary phase detector PD. If the chaotic oscillator is in a periodic state, then the process $x(\tau)$ will be synchronous with respect to $A_0 \sin \omega\tau$ and the output of the phase detector will be a constant level. In the chaotic state, the bursts of the signal $x(\tau)$ are period skips, which is detected by the phase detector. The number of pulses N counted by the counter through the detection time τ_d is fed to the threshold device. If the number of bursts N is below the threshold h', the hypothesis of the absence of a detectable signal is accepted, otherwise it is assumed to be present.

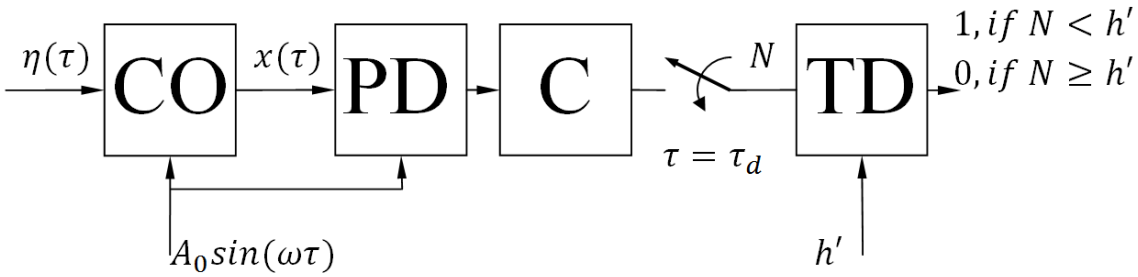


Figure 2: Detector structure diagram

The model was created in Matlab / Simulink according to the scheme presented in Fig. 3. The RandomNumber pseudo-random number generator together with the BandpassFilter filter is a source of narrowband noise with a relative bandwidth of $\gamma = 0.1$ and a center frequency ω. For each numerical experiment, 1000 launches were performed with different Seed value of the source RandomNumber. The block of the binary phase detector and counter is designated in the model as Subsystem. The time interval τ_d corresponds to $Z = 500$ periods of the sinusoidal source SineWave. The calculation was performed by the Adams method with an adaptive step. The obtained number of chaotic bursts of N was saved to a file for further statistical processing.

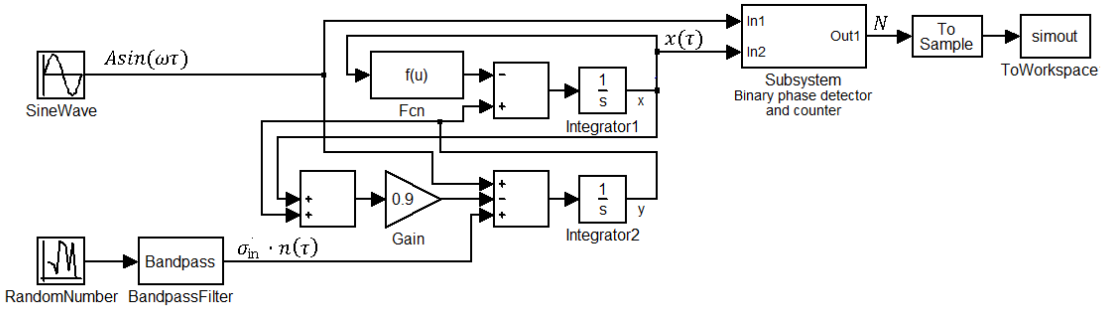


Figure 3: Numerical model of detector in Matlab / Simulink

3 Numerical experiments

The experiments investigated the statistics of the number of chaotic bursts N during the detection time at a different amplitude of a sinusoidal source A, such that $A > A_{cr}$ and different RMS of noise σ_{in}. For each numerical experiment, the mean number of bursts N_m, the standard deviation σ_N were obtained and a normalized histogram was constructed. It was found that in all experiments the mean number of bursts is less than the variance, i.e. $N_m < \sigma_N^2$. The case of overdispersion means that a model based on the Poisson distribution cannot be used. If it is necessary to set the mean and variance independently, then a model based on the negative binomial distribution can be used [1]:

$$p_{NB}(N, r, \varrho) = \binom{N + r - 1}{N} \varrho^r (1 - \varrho)^N, \tag{4}$$

where $p_{NB}(\cdot)$ is the probability mass function of the negative binomial distribution, r and ϱ are distribution parameters obtained from the expressions:

$$\varrho = \frac{N_m}{\sigma_N^2}, r = \frac{N_m}{\frac{1}{\varrho} - 1}.$$

Using the criterion χ^2 the test was performed to ensure that the experimental data corresponded to the negative binomial distribution. For this, the chi2gof() function from Matlab was used. The data on the number of degrees of freedom ν_{DF} and the significance level q_{SL} are summarized in Table 1. For most experiments, the hypothesis of belonging to a negative binomial distribution was assumed with $q_{SL} > 0.1$.

Fig. 4 - 9 show some normalized histograms (1) and corresponding theoretical dependencies (2).

Table 1: Results of numerical experiments and their verification by the criterion χ^2

σ_{in}		Amplitude,A			
		0.083	0.085	0.09	0.11
$2 \cdot 10^{-3}$	N_m	24.67	1.41	0	0
	σ_N	8.62	2.03		
	q_{SL}	0.02	0.12		
	χ^2	51.4	12.74		
	ν_{DF}	33	8		
$1 \cdot 10^{-2}$	Nm	46.2	35.03	14.9	0.041
	σ_N	10.0	9.19	6.02	0.26
	q_{SL}	0.11	0.47	0.16	0.74
	χ^2	49.57	35.8	31.78	0.1
	ν_{DF}	39	36	25	1
$2 \cdot 10^{-2}$	Nm	52.46	46.75	33.46	4.9
	σ_N	10.78	10.43	9.44	3.22
	q_{SL}	0.91	0.79	0.54	0.026
	χ^2	29.16	32.35	34.42	25
	ν_{DF}	41	40	36	14
$3 \cdot 10^{-2}$	Nm	49.15	45.72	37.8	13.88
	σ_N	10.31	10.22	9.44	5.89
	q_{SL}	0.63	0.83	0.26	0.51
	χ^2	36.45	29.51	40.87	24.16
	ν_{DF}	40	38	36	25

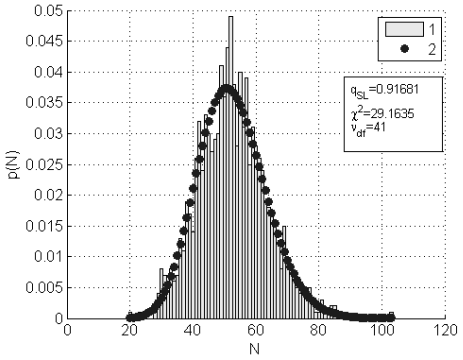


Figure 4: $A = 0.083, \sigma_{in} = 2 \cdot 10^{-2}$

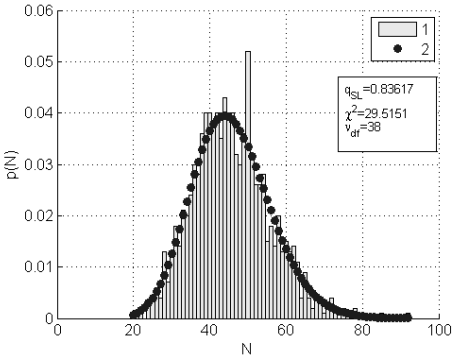


Figure 5: $A = 0.085, \sigma_{in} = 3 \cdot 10^{-2}$

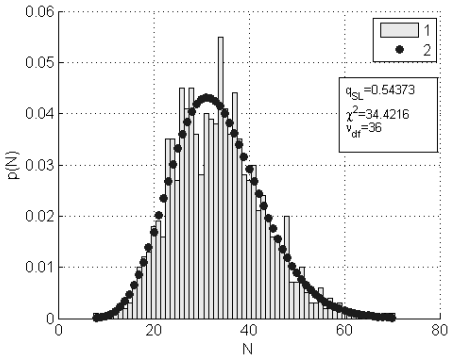


Figure 6: $A = 0.09, \sigma_{in} = 2 \cdot 10^{-2}$

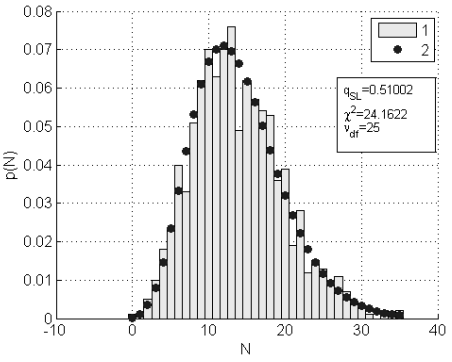


Figure 7: $A = 0.11, \sigma_{in} = 3 \cdot 10^{-2}$

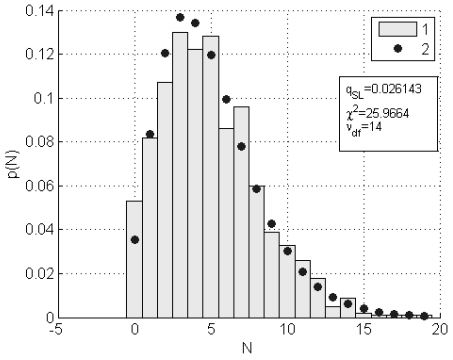


Figure 8: $A = 0.11, \sigma_{in} = 2 \cdot 10^{-2}$

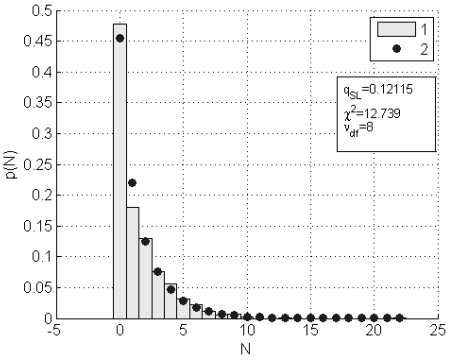


Figure 9: $A = 0.085, \sigma_{in} = 2 \cdot 10^{-3}$

Conclusions

The completed studies confirm the hypothesis of the experimental data for the number of chaotic bursts has a negative binomial distribution. Of all the numerical experiments, only the results in two cases did not show a sufficient significance level. However, it can be explained by the discrepancy in the solution of the Cauchy problem for an unstable system over large time intervals. From this it follows that for the detector with the threshold established by the Neumann-Pearson criterion, the following expression can be written:

$$h' = \Psi^{-1}(p_F, r, \varrho),$$

where $\Psi^{-1}(\cdot)$ is the inverse function of negative binomial distribution, p_F is probability of a false alarm. Accordingly, for the probability of correct detection p_D, we can write:

$$p_D = \Psi(h', r, \varrho),$$

where $\Psi(\cdot)$ is the function of the negative binomial distribution.

This work shows that the use of a chaotic oscillator to detect a signal under the background noise is possible using a statistical approach to calculating the decision threshold at different signal-to-noise ratios.

References

[1] Gschlobl S., Gzado C. (2008). Modelling Count Data with Overdispersion and Spatial Effects. *Statistical Papers*. Vol. **49**, pp. 531-552.

[2] Hramov A.E., Koronovskii A.A., Kurovskaja M.K. (2007). Length Distribution of Laminar Phases for Type-I Intermittency in the Presence of Noise. *Physical Review E*. № **76**, 026206.

[3] Jalilvand A., Fotoohabadi H. (2011). The Application of Duffing Oscillator in Weak Signal Detection. *Transactions on Electrical Engineering, Electronics, and Communications*. Vol. **9**. № 1, pp. 1-6.

[4] Jin T., Zhang H. (2011). Statistical Approach to Weak Signal Detection and Estimation Using Duffing Chaotic Oscillators. *Science China Information Sciences*. Vol. **54**. № 11, pp. 2324-2337.

[5] Lakshmanan M., Murali K. (1996). *Chaos in Nonlinear Oscillators: Controlling and Synchronization*. World Scientific Publishing, Singapore.

[6] Shi H., Li W. (2017). The Application of Chaotic Oscillator in Detecting Weak Resonant Signal of MEMS Resonator. *Review of Scientific Instruments*. № **88**, 055003.

[7] Wu A., Mwachaka S.M., Pei Y. (2018) A Novel Weak Signal Detection Method of Electromagnetic LWD Based on a Duffing Oscillator. *Journal of Sensors*. Vol. **2018**, 5847081.

Real-time multiple object tracking algorithm for adaptive traffic control systems

Mikhail A. Kovalenko[1], Natalia A. Sergeeva[2]
[1] *Siberian Federal University, Krasnoyarsk, Russia*
[2] *LLC Rd-science, Krasnoyarsk, Russia*
e-mail: mihail.kovalenko@rd-science.com
n.sergeeva@rd-science.com

Abstract

This study is dedicated to development of a tracking algorithm. The main objective for such algorithm is its mandatory implementation with minimal computing with minimal computing resources, including the provision of a data transmission channel. The requirement is dictated by the need to obtain information about the ratio of vehicles passing through the intersection in real time. The data is necessary for adaptive correction of the time shifts for coordination plans of traffic lights. The development of the proposed method assumes the classification of the received tracks for the recognition of emergency situations.

Keywords: real-time tracking, object recognition, convolutional neural network, distance metrics, machine learning, computer vision, adaptive traffic control systems.

Introduction

At these days video processing technologies are widely used for solving various kinds of recognition, control and optimization problems. In particular, to optimize the lengths of the traffic light phases it is necessary to know the ratio of the traffic flows at the intersection. Knowing the number of cars traveling in different directions will allow to make a decision about changing the traffic pattern, the phase sequence, adjusting the length of each phase or calculating the offset of the traffic light cycle of one intersection relative to its neighbors.

Thereby the task of constructing, recognizing and classifying the tracks of moving objects becomes important.

The solution of the tracking problem includes the recognition stage and tracking stage. There are two most common approaches: object tracking based on finding key points of an object, a significant number of these methods are implemented in the OpenCV library [1], and tracking algorithms based on object classification and localization using convolutional neural networks [2].

The authors of this study used the second aproach, the main focus of the research is on the development of a track constructing algorithm. Object recognition and localization was implemented using the retrained model of YOLO v.3.0 [3]. Tracking methods implemented in the OpenCV library provide high stability and accuracy of tracking moving objects. However, this approach does not imply automatic selection of the tracked objects, as well as the classification of these objects. And provided

with the dozens of tracked objects the processing speed of a frame becomes on par with the processing speed of the convolutional neural network.

Therefore, it was decided to use the convolutional neural network with the subsequent processing of the output data to form the trajectories of objects. This approach gives more necessary information such as the class of an object. Also it has a constant processing speed that does not increase with the number of objects in the frame.

A significant advantage of this algorithm is its ability to obtain the desired results in real time at the installation site of the camera. Transferring a video stream to a server for processing was considered inefficient and costly. The instability of the transmission of high-quality images for processing and a significant number of data collection points for the adaptive traffic control system requires achieving the high accuracy of data processing results at the low cost of computing resources.

1 Image processing stage

At the stage of detection, the neural network brings the original image to the set square size and divides it into square blocks based on the hyper-parameter S. Their number is calculated as $S \times S$. Each block at the detection stage makes an estimation about objects whose centers are located inside the block. The estimation includes the coordinates of the object center inside the block and its dimensions. The number of estimations that is produced by each block is set by the hyper-parameter Q.

Thus, in the course of image processing, the output from each block is the following matrix:

$$B = \begin{pmatrix} x_1 & y_1 & h_1 & w_1 & c_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_q & y_q & h_q & w_q & c_q \end{pmatrix}, \tag{1}$$

where x_i, y_i – are the coordinates of the i-th object (in pixels), h_i, w_i – are the height and width of this object (in pixels), c_i – is the confidence level.

Each block also makes a class prediction based on the features that were extracted during the convolution process. Depending on the number of classes that were initialized during the training process of the network, the number of output parameters will also differ. Thus, the output for each image block is the vector $\vec{p}$:

$$\vec{p} = (p_1, p_2, \ldots, p_k), \tag{2}$$

where p_i – is the probability of the i-th class in this block, k – the number of all classes.

Next, a decision is made about the presence of an object inside the block or its absence. The decision is made according to the following rule:

$$(\exists i \in B) c_i > t, \tag{3}$$

where c_i – confidence level of the i-th estimation, t – the specified threshold for the confidence level.

That is, if there is at least one estimation with the confidence level exceeding a given threshold t – a decision is made about the presence of an object inside the block. Its coordinates and dimensions are set as:

$$\begin{aligned} d_j = d_i, \quad d_i \in B_j, \\ c_i = max(c_k), \quad c_k \in B_j, \end{aligned} \tag{4}$$

where d_j – is the dimension (x, y, w or h) of the j-th object, c_k – confidence level of the k-th estimation of the block, B_j – is the B matrix of the j-th object.

The class of an object is set as:

$$a_j = h, \quad p_h = max(\vec{p_j}) = max(p_j1, p_j2, \ldots, p_jk), \quad h = \overline{1,k}, \tag{5}$$

where a_j — class of the j-th object, p_j – probability vector corresponding to the j-th object.

Therefore, each image block generates either one object or none. The final set of objects forms the output matrix:

$$V = \begin{pmatrix} x_1 & y_1 & h_1 & w_1 & a_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_m & y_m & h_m & w_m & a_m \end{pmatrix}, \tag{6}$$

where x_i, y_i – are the coordinates of the i-th object (in pixels), h_i, w_i – are the height and width of this object (in pixels), a_i – is the number corresponding to the class of the object, m - is the resulting number of objects on the frame.

The error rate of the classification module is estimated to be 93.8% [3].

2 Tracking stage

Processing video frame by frame we get a constantly updated sequence of recognized objects V. Further processing of these objects in order to track each object thus turns into the association problem for two consecutive frames V_i and V_{i-1} (Figure 1).

List of objects is not permanent. In the considered task of detecting vehicles new objects constantly appear on the frame and old objects disappear with the same rate. Moreover, detection using a neural network still has a certain error rate – the object may not be recognized on one of the frames, but reappear on the next one.

In order to provide a constantly up-to-date list of objects, the V_i, the i-th list of objects, is compared not with V_{i-1}, but with a constantly updated list of objects V'. Objects are deleted from V' only after the number of consecutive frames on which the object is absent will overcome the threshold l.

The general detection algorithm is presented below.

1. Received a list of objects V_i.
2. Association between objects V_i and V'.
3. Adding new objects from V_i to V'.

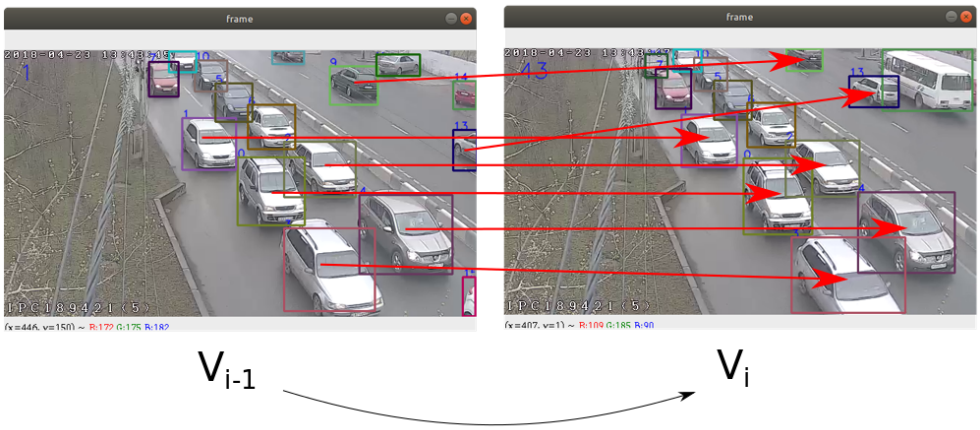


Figure 1: Object association between two consecutive frames

4. Removing old objects from V'.

The association between objects V_i and V' is made by analyzing the centroids of each object. For each object from V_i, there is an object closest to it from V'. To estimate the distances between objects, the Euclidean metric for two-dimensional space is used:

$$\begin{aligned} dist_j = min(d(v_j, v'_k)) = min\{d(v_j, v'_1), d(v_j, v'_2), \ldots, d(v_j, v'_m)\} = \\ = min\left\{\sqrt{(x_j - x_1)^2 + (y_j - y_1)^2}, \ldots, \sqrt{(x_j - x_m)^2 + (y_j - y_m)^2}\right\}, \end{aligned} \tag{7}$$

where $dist_j$ – is the minimal distance from v_j, v_j – the j-th object from V_i, v'_k – the k-th object from V' with $k = \overline{1, m}$, m - the number of objects on the frame.

Then all of the V_i objects are sorted by their corresponding $dist_j$ values. Each object, starting from the objects with the smallest $dist_j$, is associated with the object from V' closest to it. Association is made only for those objects whose shortest distance does not exceed the threshold calculated as:

$$\Delta = l \times \sqrt{w^2 + h^2}, \tag{8}$$

where l – the number of frames on which the object was absent, w – found width of the object (in pixels), h – its height (in pixels).

Each object from V' to which no object from V_i has been assigned is marked as missing on the frame. If the object is absent on 5 frames in a row, it is declared disappeared and is removed from V'.

All objects from V_i, for which there was no match, are added to V' as new objects that were not detected earlier.

The results of the tracking system can be seen in Figure 2.

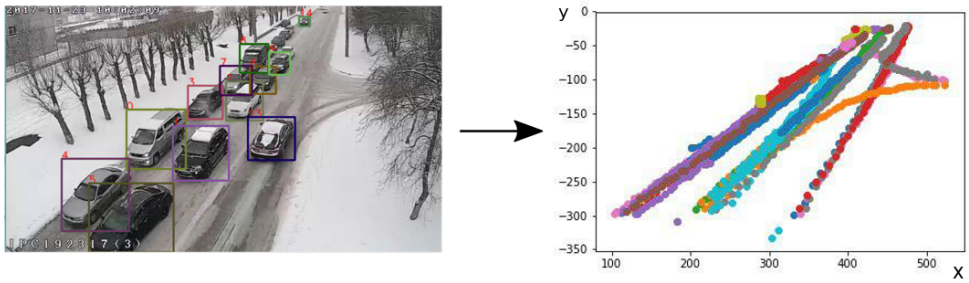


Figure 2: Results of the tracking algorithm

3 Track processing

The tracks of objects, after they leave the frame are processed for further work.

The class of the same object can vary on different frames due to the probability of error during the recognition stage. To establish the true class of an object, a metric was introduced that assesses the probability that an object belongs to a particular class. When processing each frame, the neural network makes its contribution to one of the three indicators, each of which corresponds to one of the considered classes.

For the class that was detected for the given object on the i-th frame, this indicator is calculated as:

$$g_i = g_{i-1} + \frac{G}{\sqrt{w^2 + h^2}}, \tag{9}$$

where g_i – indicator for the given object on the i-th frame, w – detected width of an object on the i-th frame (in pixels), h – its detected height (in pixels), G – special parameter.

For all the other classes on the i-th frame, the indicator g is transferred from the previous frame:

$$g_i = g_{i-1}. \tag{10}$$

When it is necessary to determine the class of an object, the following decision rule is then applied:

$$a_T = j, \quad g_j = max(g_i), \quad i \in C, \tag{11}$$

where a_T – is the true class of an object, g_j – the indicator of an object for the j-th class, C – the whole set of classes established in the system.

The parameter G is relative and is used exclusively for comparison between classes. The parameter G can be set to different values depending on the class in order to give priority to one class over the others. At the moment setting the parameters for every class remains entirely up to the analyst. In the future, it is possible to present the settings of this parameter as an optimization problem.

Collected tracks are compared with pre-marked areas for car counting. The collected data is divided by classes and the traffic directions and after that sent through the communication channels to the target system.

The target system is an automated traffic management system, namely, a module that allows modeling and simulation of various road junctions.

The collected information on the number of cars makes it possible to simulate traffic at the considered intersection, most accurately reflecting the actual situation for a given period of time. The analysis of the historical data collected using the developed tracking system allows to predict the situation at the intersection for the next period of time.

Substituting the predicted intensity data into the modeling system, it becomes possible to automatically select the optimal parameters for the traffic light. An example of such a system can be seen in Figure 3.

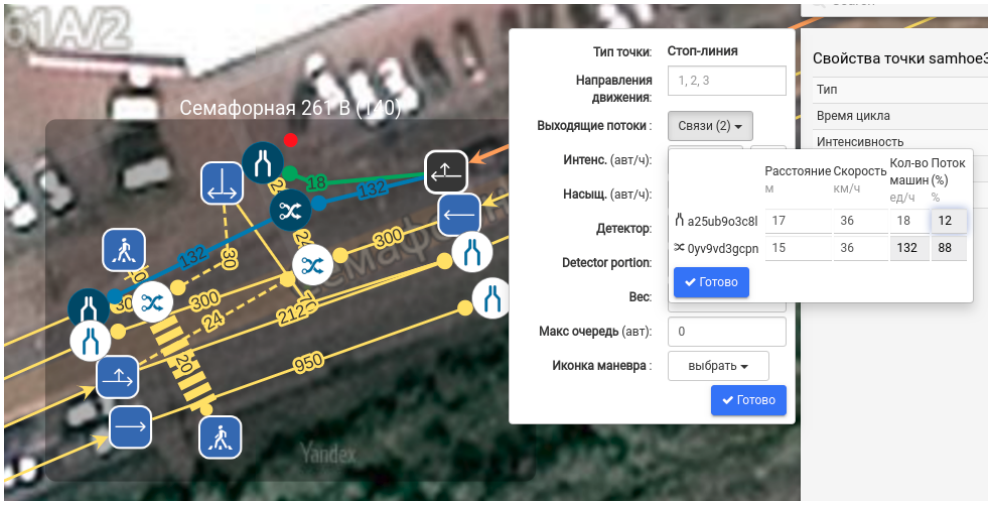


Figure 3: The simulation module of the automated traffic management system

Information divided by different directions gives an idea of the ratio of those traffic flows. Knowing the number of passing vehicles of different classes allows to estimate the speed of the transport (since public transport and trucks move slower than other cars).

Conclusions

The proposed tracking has shown a high accuracy of track construction under various conditions on the transport network. The algorithm takes into account the possible absence of the object in the surrounding frames. The frame processing frequency of the system reaches 25 frames/s, which makes it possible to talk about real-time multiple object tracking. Further development of the work is aimed at clustering the tracks to assess the abnormal or emergency situations on a section of the road. In case of the abnormal traffic or trajectories at the observable intersection, correction

of the coordination plans of several traffic lights is required to reduce the negative impact of the accident.

References

[1] OpenCV: Tracking API - OpenCV Documentation [Online]. Available: https://docs.opencv.org/3.4/d9/df8/group__tracking.html

[2] Agarwal, A., Suryavanshi, S. (2017). *Real-Time* Multiple Object Tracking (MOT) for Autonomous Navigation*. Technical report.

[3] Redmon, J., Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.

Disproportion of Russian Regions development in the sphere of population provision with food of own production

Vladimir Glinskiy, Lyudmila Serga, Yuliya Ismaiylova and Michael Alekseev
Novosibirsk State University of Economics and Management, Novosibirsk, Russian Federation e-mail: s444@ngs.ru

Abstract

A component of national security is food security, which includes providing the population with food in sufficient quantity and good quality. Russia adopted a strategy towards its own food production. However, not every region has natural and climatic conditions and resource potential for food production, and the vastness of the territory of Russia creates problems in the regularity of interregional food supplies. This leads to disproportion of territorial development in the sphere of providing the population with own food production. The article presents the results of a study of regional unevenness and differentiation of the development of the food industry since 2010 and the problems of eliminating the disproportion of the population's food security. The methodology is based on the use of statistical tools: methods of classification, evaluation of regional developmental unevenness, calculation of the coefficients of differentiation and concentration of production.

Keywords: sustainable development; Russian regions development; food security; disproportion; differentiation.

Introduction

The second goal of the Program of Transforming Our World, adopted by the UN is to eradicate hunger, achieve food security and improve nutrition. The food and agriculture sector offers key solutions for development, and is central for hunger and poverty eradication. If done right, agriculture, forestry and fisheries can provide nutritious food for all and generate decent incomes, while supporting people-centred rural development and protecting the environment. [1]

Food security is the providing the population with food in sufficient quantity and good quality. [2–12] There are two main sources of providing the population with food: own production and import. Since Russia is under sanctions from the United States and the European Union, it is forced to adopt a strategy for its own food production. Therefore, each region of Russia is trying to develop the food industry to provide for own residents with natural foods. [13–19] However, not every region has natural and climatic conditions and resource potential for food production. As a result, firstly, there is uneven development of the food industry in the regions of the country [20–24]; secondly, the vastness of the territory of Russia creates problems in the regularity of interregional supplies of agricultural raw materials. This leads to

disproportion of territorial development in providing the population with food products of their own production. [25–33] The food industry sector includes more than 40 sub-industries, however to address the research problems, the authors consider the food industry in all regions of the Russian Federation without division into product types. This will make it possible to analyze and draw general conclusions about the level of provision of the population with food of own production.

1 Estimation of regions differentiation level

The uneven development of the food industry can be viewed from two perspectives: the unevenness of production volumes by regions of the Russian Federation and the uneven supply of food for the population.

First, the authors propose to consider the disparity in the development of regions in the provision of food for the population as the uneven development of the food industry. To determine the degree of differentiation in the development of the food industry, the Lorenz curve was constructed in the regions for the volume of food products of own production for 80 regions of the Russian Federation in 2011, 2013 and 2015. [34–36]

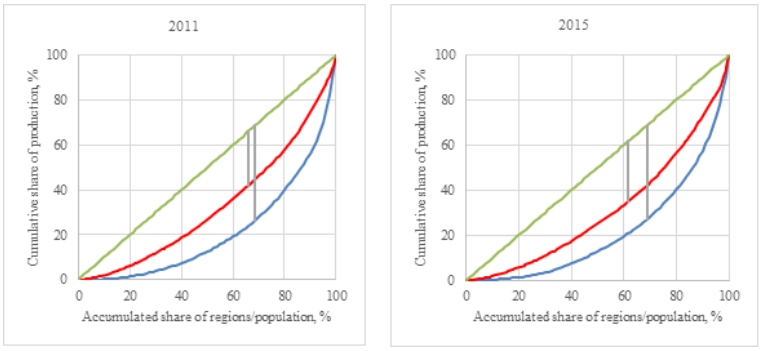


Figure 1: Fig. 1. (a) The Lorenz curve for 2011; (b) The Lorenz curve for 2015

The Lorentz curve (Fig. 1) was constructed in two ways: in the first case, the regions of the Russian Federation are indivisible units (on the lines a and b the lines are marked in blue, and the accumulated share of the regions was plotted along the X axis), in the second case, the adjustment is made by the population in each object (the red line in the charts, and the accumulated share of the population was plotted along the X axis). On the Y-axis, in both cases, the cumulative production share was presented. In the second case, the deviation of the curve from the uniform distribution is less than in the first, that is, the degree of differentiation of the regions taking into account the structure below.

In order to trace the degree of inequality in dynamics, the Gini coefficients were calculated for Lorenz curves of the second type (taking into account the structure of the population by regions) from the formula (1).

$$K_G = 1 - 2\sum_{i=1}^{n} x_i \cdot cumy_i + \sum_{i=1}^{n} x_i \cdot y_i \quad (1)$$

Where x_i – the share of the population in the i-th region; y_i – the share of food production in the i-th region; n – number of regions

The Gini coefficient for the Russian Federation in 2011, 2013 and 2015 was 0.349, 0.347 and 0.376, respectively. Despite the low Gini coefficients, it can be concluded that before the introduction of sanctions, the differentiation of the food industry in the regions was declining, and now it is growing.

For the same time periods, the Hoover indices, the so-called Robin Hood indices – the maximum deviation of the Lorentz curve from the equality line – were calculated from the formula (2):

$$I_G = max_i(i - F(x_i)) \quad (2)$$

Where $F(x_i)$ – is the point on the Lorentz curve.

For the Lorentz curves of the second type, the Hoover index was 24.3, 23.9, and 26.7. Thus, the Hoover index confirms the growth of differentiation of regions in terms of their own food production. Therefore, in order to achieve territorial uniformity of production, it is necessary to develop the food industry in lagging regions more rapidly.

2 Typology of regions

For a more detailed analysis of the degree of differentiation in the development of the food industry, a typological grouping of the regions of the Russian Federation was carried out using the adapted BCG matrix, which makes it possible to compare the positions of the Russian Federation regions based on a combination of the rates of growth in production value (the Y-axis) and the relative share of production value (the X-axis). The regions of the Russian Federation are represented in the diagram below (Fig. 2) in the form of circles whose radius is proportional to the per capita production value in each region. To determine the positions that the regions occupy, the Y-axis was divided into two categories - low and high growth rates (the critical point is the rate of growth in the Russian Federation), the X axis - low and high relative shares of production values of individual regions as compared to the average Russian share (critical point is 1). This allowed to obtain 4 types of regions in terms of providing the population with own products of food industry enterprises: donors, self-sufficient, prospective and recipients:

- Recipients are characterized by low rates of growth in production value and a low relative share of production. The recipient region is unable to provide its population with the necessary food and is forced to import it.
- Self-sufficient regions are regions with low rates of growth and a high relative share of production.

- Perspective – high rates of growth, low relative share of production. Perspective region is one of the most unstable position, which requires additional research. Without additional actions, prospective regions can easily move to the category of recipient regions
- Donors are characterized by high growth rates and a high relative share of production. At the expense of the donor regions, food products are provided to the population of the recipient regions.

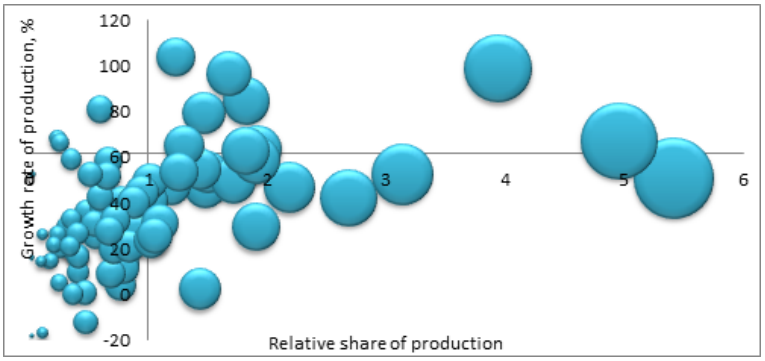


Figure 2: Adapted BCG - matrix by regions of the Russian Federation

Fig. 2 shows the regions of the Russian Federation according to data for 2015, the growth rates were calculated in comparison with 2013. The result of the distribution of regions by groups is presented in Table 1.

Table 1: Types of Russian Federation regions by production volumes

	2013/2011	2015/2013
Donors	16	24 (12 sustainable)
Self-sufficient	17	5 (4 sustainable)
Perspective	23	15 (7 sustainable)
Recipients	24	36 (15 sustainable)

Territorial distribution of regions into 4 types is shown in Fig. 3. Thus, we can conclude that in the development of the food industry there are both positive and negative trends. The positive changes include the growing number of donor regions, 12 sustainable regions in the group of donor regions, and the growth is provided mainly by self-sufficient regions, which not only retain a relative share of production, but also increase production rates. Also, positive changes include the fact that in the category of perspective regions in 2015, there were 8 former recipient regions. A negative trend is the growth of differentiation of regions. Moreover, an increase in the number of recipient regions was inevitable due to a change in the foreign

policy situation and the imposition of sanctions, which entailed the need for financial investments in the food industry, as well as the adoption of timely and competent management decisions.

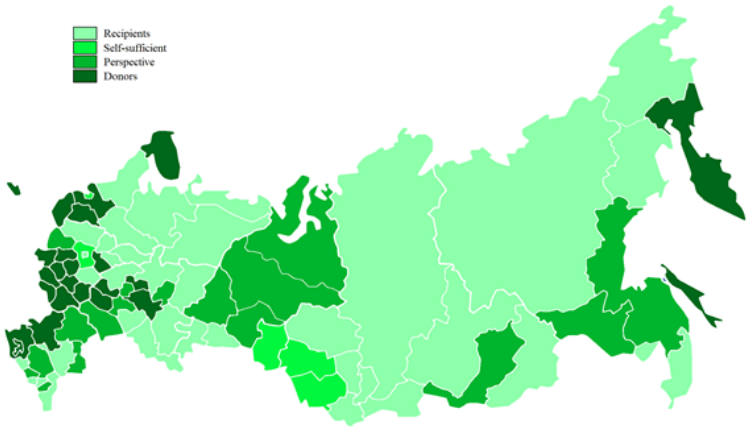


Figure 3: Cartogram of the distribution of Russian regions by the results of a typological grouping

3 Typological grouping of regions in terms of food security for the population

Further, the authors propose to analyze the provision of the population with food. For this purpose, the typology of subjects of the Russian Federation in terms of production per capita was implemented. The regions with a high level were those regions of the Russian Federation whose average per capita production level is higher than the middle for the Russian Federation by an amount greater than the standard deviation to regions with a low level - those entities whose per capita production level is significantly less than the average deviation. The results of typologization are presented in Table 2.

Table 2: Distribution of the Russian regions by the provision of the population with food (production per capita).

The provision of the population with food	2011	2013	2015
High level	7	6	3
Middle level	65	66	71
Low level	8	8	6

Thus, it can be traced that the degree of differentiation decreases over time – an ever-smaller number of regions fall into the tails of distribution.

For the analysis of the degree of variability of average per capita production values, the coefficient of variation was calculated, which amounted to 1.22, 1.23 and 1.03 in 2011, 2013 and 2015, respectively, i.e. over time, the variation decreases, which is due to an increase in the number of donor regions and domestic imports of own-produced food.

Conclusions

Thus, the results of the study indicate that during the period under review, the disparities between the Russian Regions Development in the Sphere of Population provision with the Food of Own Production have grown. This was due to the regions with high level of food production per capita. But simultaneously there is a shift towards the concentration of regions around the average level. This is a positive factor, as it stimulates the growth of own food production in the Russian Federation.

References

[1] https://www.un.org/sustainabledevelopment/ [Electronic source] (11.04.2019).

[2] The Doctrine of Food Security of the Russian Federation: Decree of the President of the Russian Federation of January 30, 2010 No. **120**.

[3] A. Shamaii, M. Omidvari, F.N. Lotfi (2017). Health, safety and environmental unit performance assessment model under uncertainty (case study: steel industry). *Environmental Monitoring and Assessment*. 189 (1), 42.

[4] A.N. Chekavinskiy, R.Yu. Selimenkov (2014). Modelling of food security in the region. *Economic and social changes: facts, trends, forecast*. No. **4**(34), pp. 226-235.

[5] V. Glinskiy, L. Serga, M. Khvan, K. Zaykov (2016). Fuzzy Neural Networks in the Assessment of Environmental Safety. *Procedia CIRP*.No **40**, pp. 615-619.

[6] V. Glinskiy, L. Serga, M. Khvan, K. Zaykov (2017). A Spatio-dynamic Modelling of Environmental Safety of the Russian Federation Regions. *Procedia Manufacturing*. No **8**, pp. 315-322.

[7] I. Kirilchuk, A. Barkov, L. Shul'ga (2017). Assessment and GIS analysis of the human health risk from negative emissions into the air. *International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management, SGEM*. No **17**(51), pp. 117-124.

[8] Y. Shi, J. Li, M. Xie (2018). Evaluation of the ecological sensitivity and security of tidal flats in Shanghai. *Ecological Indicators*. No **85**, pp. 729-741.

[9] I. Kirilchuk, V. Yushin, V. Protasov (2016). Improvement of the models and algorithms of social and hygienic monitoring in the system of air protection activities. *International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management, SGEM*. No **2**, pp. 455-462.

[10] Y. Bezsonov, V. Andreev (2016). Justification and formalization of approach to regional environmental safety evaluation. *Eastern European Journal of Enterprise Technologies*. No **2**(10), pp. 9-18.

[11] A. Larionov, Y. Larionova (2017). Economic-legal Securing of Integrated Territorial Development for Purposes of Housing Construction in the Moscow Region. *MATEC Web of Conferences*. No **106**, 08032. DOI: 10.1051/matecconf/20171060

[12] M.F. Tyapkina, N.A. Antipyeva (2010). Food security of the region is an integral part of the national security. *Bulletin of IrGSKhA*. No **39**, pp. 88-100.

[13] The concept of long-term socio-economic development of the Russian Federation for the period up to 2020: Decree of the Government of the Russian Federation of November 17 (2008) No. **1662**.

[14] The fundamentals of the state policy of the Russian Federation in the field of healthy nutrition for the population until 2020: the order of the Government of the Russian Federation of October 25 (2010) No. **1873**.

[15] W. Liefert (2002). Comparative advantage in Russian Agriculture. *Agriculture Economics*. No **84**(3), pp. 762-767.

[16] D. Shields (2015). Farm Safety Net Programs: Background and Issues. *Congressional Research Service Report*. August 21, URL: http://www.fas.org/sgp/crs/misc/R43758.pdf

[17] O.V. Borisova (2006). Food industry as a factor of sustainable development of rural areas. *Nikonov readings*. No **11**, pp. 263-265.

[18] V.D. Goncharov (2009). Food industry in the system of ensuring food security of Russia. *Nikonov readings*. No **14**, pp. 35-37.

[19] V.V. Orlov, Yu.V. Klokov (2009). Scientifically grounded regulation of the region's food industry. *Food industry*. No **8**, pp. 20-21.

[20] J.G. Williamson (1965). Regional Inequality and the Process of National Development: A Description of Patterns. *Economic Development and Cultural Change*. 4, Vol **13**, part 2, pp. 1-84.

[21] S.P. Jenkins, P. van Kerm (2009). The Measurement of Economic Inequality. *The Oxford Handbook of Economic Inequality, Oxford: Oxford University Pressn*. pp. 40-67.

[22] F.A. Cowell (2009). Measurement of Inequality. *Handbook of Income Distribution*. Vol. **1**, Amsterdam: North Holland, pp. 87-166.

[23] A.B. Atkinson (1970). On the Measurement of Inequality. *Journal of Economic Theory*. No **2**(3), pp. 244-263.

[24] V. Glinskiy, L. Serga, A. Novikov, G. Litvintseva, A. Bulkina (2017). Investigation of Correlation between the Regions Sustainability and Territorial Differentiation. *Procedia Manufacturing*. No **8**, pp. 323-329.

[25] S. Beliakov, A. Kapustkina (2016). Analysis of Performance Indicators of Functioning of Territories with Special Economic Status in the Russian Federation. *Procedia Engineering*. No **165**, pp. 1424-1429.

[26] P. Graboviy (2016). Methods of Motivation Improvement and Effectiveness Increase on the Example of Construction Industry Enterprises. *Procedia Engineering*. No **165**, pp. 1520-1528.

[27] D.V. Zavyalov, O.V. Saginova, N.V. Zavyalova (2017). The concept of managing the agro-industrial cluster development. *Journal of Environmental Management and Tourism*. No **8**(7), pp. 1427-1441.

[28] Expert.ru [Electronic source] (11.04.2019).

[29] A.A. Gladysheva, T.A. Ratnikova (2014). The role of heterogeneity and mutual influence of Russian regions in the distribution of foreign direct investment in the food industry. *Economic Journal of the HSE*. No **2**.

[30] T. Buccellato, F. Santangelo. (2009). *Foreign Direct Investments Distribution in the Russian Federation: Do Spatial Effects Matter?*. Economics Working Papers 99, London: Center for the Study of Economic and Social Change in Europe, SSEES, UCL.

[31] M. Tracy. (1995). *Agriculture and Food in the Economy of Developed Countries: An Introduction to Theory, Practice, and Politics, Transl. with English*. St. Petersburg: Economic School.

[32] The State of Food Insecurity in the World [Electronic resource], FAO, IFAD, WFP/-2015/, URL: http://www.fao.org/3/ai4646e.pdf (11.04.2019)

[33] V. Glinskiy, L. Serga, K. Zaykov (2017). Identification Method of the Russian Federation Arctic Zone Regions Statistical Aggregate as the Object of Strategy Development and a Source of Sustainable Growth. *Procedia Manufacturing*. No **8**, pp. 308-314.

[34] www.gks.ru [Electronic source] (11.04.2019)

[35] Single interdepartmental information and statistical system (EMISS) [Electronic resource], URL: http://fedstat.ru (11.04.2019)

[36] Central statistical database given [Electronic resource], URL: http://cbsd.gks.ru (11.04.2019)

A nonparametric approach for estimating the set of solutions of random linear programming

Boris S. Dobronets and Olga A. Popova
Siberian Federal University, Krasnoyarsk, Russia
e-mail: BDobronets@yandex.ru, OlgaArc@yandex.ru

Abstract

The article discusses a new approach to linear optimization with random input data. We use computational probabilistic analysis to construct probabilistic extensions and set of solution of systems of linear algebraic equations with random parameters. Estimates of the probability density function of the objective function of the random linear programming problem are constructed.

Keywords: Computational probabilistic analysis, probabilistic extensions, random systems of linear algebraic equations, linear programming, random input data.

Introduction

The study of many practical problems, including the problem of decision-making, requires the implementation of the optimization approach. The effectiveness of the solutions depends on several factors. These factors primarily include the data needed to describe and solve the problem. One of the important factors that should be considered when solving such problems is uncertainty of input data.

The paper deals with the computational probabilistic approach to solving optimization problems with random inputs. Using methods of mathematical programming, we obtain optimal solutions that depend on these parameters. When probability densities of input parameters are known, it is possible to construct a joint probability density function for the optimal solutions.

In most of stochastic programming algorithms the operator of mathematical expectation is used and averaging procedures are performed [10]. Note that in stochastic programming the optimal solution is a fixed point. In the case of random programming, a joint probability density function is constructed on the set of solutions of the optimization problem.

To this end, nowadays mathematical tools for uncertain programming are developed. Uncertain programming is the theoretical basis for solving optimization problems for various uncertainty conditions. In [7] three main types of uncertainty, namely, randomness, fuzziness, and imprecision, are recognized. Since an interval number can be considered as a special case of an imprecise quantity, interval analysis, interval arithmetic, and interval programming fall into imprecise programming [5].

This work continues the research started in [2]. In the present paper, special attention is paid to the construction of probability density functions of the solution sets of systems of linear algebraic equations with random coefficients. In contrast to

the works [4, 9], the considered approach does not require restrictions on the type of probability density functions of the input data.

1 Formulation of the problem and background

Let us formulate the problem of *random programming* as follows:

$$\min f(x, \xi), \tag{1}$$

subject to (s.t.)

$$g_i(x, \xi) \leq 0, \quad i = 1, ..., m, \tag{2}$$

where x is the solution vector, ξ is the vector of parameters, $f(x, \xi)$ is the objective function, $g_i(x, \xi)$ are constraint functions.

As for ξ, we know that this is a random vector. The randomness of random vector involved in the above statements is formalized through a probability space, $(\Omega, \mathcal{F}, P)$, where, Ω, $\mathcal{F}$, P are the set of random events, the σ – algebra of the subsets of Ω and the applicable probability measure, respectively. The elements of this probability space appear as parameters in the input random vector $\xi(\omega), \omega \in \Omega$.

Vector x^* is the solution of problem (1) – (2), if

$$f(x^*, \xi(\omega)) = \inf_U f(x, \xi(\omega)),$$

where

$$U = \{x | g_i(x, \xi(\omega)) \leq 0, \quad \omega \in \Omega, i = 1, ..., m\}.$$

The solution set of (1)–(2) is defined as follows

$$\mathcal{X} = \{x | \min f(x, \xi(\omega)), g_i(x, \xi) \leq 0, \quad i = 1, ..., m, \omega \in \Omega\}.$$

Note that x^* is a random vector, so in contrast to the deterministic problem, for x^* it is necessary to determine the probability density function for each component of x_i^*.

When both objective function and constraint functions are linear functions, the problem is called a problem of linear programming. Otherwise, the problem is called a problem of nonlinear programming.

For example the problem of linear programming with random data is formulated as follows:

$$\min(c(\omega), x), \tag{3}$$

s.t.

$$A(\omega)x = b(\omega), x \geq 0, \omega \in \Omega. \tag{4}$$

The vector x^* is the solution of problem (3) – (4) provided that

$$(c(\omega), x^*(\omega)) = \inf_{U(\omega)} (c(\omega), x),$$

where

$$U(\omega) = \{x | A(\omega)x = b(\omega), x \geq 0, \omega \in \Omega\}.$$

The solution set of (3) – (4) is

$$\mathcal{X} = \{x | \min(c(\omega), x), A(\omega)x = b(\omega), x \geq 0, \omega \in \Omega\}.$$

The probability density functions of random variables x, y, z will be denoted bold font $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$.

We will be interested in estimating the probability density function of the objective function $(c(\omega), x^*(\omega)), \omega \in \Omega$. For these purposes we will use Computational Probabilistic Analysis (CPA).

2 Elements of Computational Probabilistic Analysis

The basis of Computational Probabilistic Analysis (CPA) is numerical operations on probability density functions of the random values. These are operations "$+$", "$-$", "$\cdot$", "$/$", "$\uparrow$", "max", "min", as well as binary relations "$\leq$", "$\geq$" and some others. The numerical operations of the piecewise polynomial function arithmetic constitute the major component of CPA. The use of CPA for these problems is more effective than the Monte Carlo method in a thousand times.

Using the arithmetic of probability density functions and probabilistic extensions, we can construct numerical methods that enable us solving systems of linear and nonlinear algebraic equations with random parameter [1].

We will use piecewise polynomial models to represent probability density functions.

One of the most important problems that CPA deals with is to construct probability density functions of random variables. Let z be a function $f(x_1, \ldots, x_n)$, where $(x_1, \ldots, x_n)$ is random vector with joint probability density $p(x_1, \ldots, x_n)$.

Definition 1. *By probabilistic extension $\boldsymbol{f}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ of the function f, we mean a probability density function $\boldsymbol{z}$ of the random variable z*

$$\boldsymbol{z} = \boldsymbol{f}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n).$$

Definition 2. *Support of the probability density functions $\boldsymbol{f}$ will be called the set*

$$supp(\boldsymbol{f}) = \{x | \boldsymbol{f}(x) > 0\}.$$

One possible way to estimate the probability density $\boldsymbol{z}$ of a random variable z

$$z = f(x_1, \ldots, x_n). \tag{5}$$

is the Monte Carlo method [8]. For these purposes a random vector $(x_1^i, \ldots, x_n^i)$ with joint probability density function $p(x_1, \ldots, x_n)$ is generated. Further we are calculated $z^i = f(x_1^i, \ldots, x_n^i), i = 1, ..., N$. Using the histogram method for z^i, we can construct an estimate of the probability density function $\boldsymbol{z}$.

Theorem 1 ([3])**.** *Let $\boldsymbol{f}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ be probabilistic extensions of function $f(x_1, x_2, \ldots, x_n)$ and for each real t function $\boldsymbol{f}(t, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ be probabilistic extensions of the function $f(t, x_2, \ldots, x_n)$. Then*

$$\boldsymbol{f}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) = \int_{supp(\boldsymbol{x}_1)} \boldsymbol{x}_1(t) \boldsymbol{f}(t, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) dt. \tag{6}$$

Corollary 1. *Theorem 1 infers the possibility of recursive computations for the general form of probability extensions and reduction to the calculation of the one-dimensional case.*

Let $f(x_1, \ldots, x_n)$ be a rational function. To construct a probabilistic extension $\boldsymbol{z}$, we replaced the arithmetic operation by the numerical operation, while the variables $x_1, x_2, \ldots, x_n$ are replaced by piecewise polynomial functions of their possible values. It makes sense to call the resulting piecewise polynomial functions of $\boldsymbol{z}$ as *natural probabilistic extension* [1].

Theorem 2 ([1]). *Let $x_1, \ldots, x_n$ be independent random variables. If $f(x_1, \ldots, x_n)$ is a rational expression where each variable x_i occurs not more than once, then the natural probabilistic extension approximates a probabilistic extension.*

3 Systems of linear algebraic equations

Let us consider solution of a system of linear algebraic equations

$$Ax = b, \tag{7}$$

where $A = (a_{ij})$ a random matrix and $b = (b_i)$ a random right-hand side vector respectively. Suppose that the random matrix A and the vector b have independent components with probability densities $\boldsymbol{A} = (\boldsymbol{a}_{ij})$, $\boldsymbol{b} = (\boldsymbol{b}_i)$ respectively and

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_{11} & \boldsymbol{a}_{12} & \ldots & \boldsymbol{a}_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_{n1} & \boldsymbol{a}_{n2} & \ldots & \boldsymbol{a}_{nn} \end{pmatrix}.$$

The support of the solution set can be represented as follows [2]

$$\mathcal{X} = \{x | Ax = b, A \in \mathrm{supp}(\boldsymbol{A}), b \in \mathrm{supp}(\boldsymbol{b})\}.$$

Construct the probabilistic extension of the solution vector $\boldsymbol{x}_1(\boldsymbol{A}, \boldsymbol{b})$

$$\boldsymbol{x}_1(\boldsymbol{A}, \boldsymbol{b}) = \frac{\begin{vmatrix} \boldsymbol{b}_1 & \boldsymbol{a}_{12} & \ldots & \boldsymbol{a}_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{b}_n & \boldsymbol{a}_{n2} & \ldots & \boldsymbol{a}_{nn} \end{vmatrix}}{\begin{vmatrix} \boldsymbol{a}_{11} & \boldsymbol{a}_{12} & \ldots & \boldsymbol{a}_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_{n1} & \boldsymbol{a}_{n2} & \ldots & \boldsymbol{a}_{nn} \end{vmatrix}}$$

or

$$\boldsymbol{x}_1(\boldsymbol{A}, \boldsymbol{b})(\xi) = \int \boldsymbol{a}_{12}(t_{12}) \ldots \boldsymbol{a}_{nn}(t_{nn}) \frac{\sum \boldsymbol{b}_i \Delta_i(t_{12}, \ldots, t_{nn})}{\sum \boldsymbol{a}_{1i} \Delta_i(t_{12}, \ldots, t_{nn})}(\xi) dt_{12} \ldots dt_{nn}, \tag{8}$$

where $\Delta_i(t_{12}, \ldots, t_{nn}) \in R$ are minors from the Cramer method for solving SLAE, $t_{ij} \in \mathrm{supp}(\boldsymbol{a}_{ij})$. The expression

$$\left(\frac{\sum \boldsymbol{b}_i \Delta_i(t_{12}, \ldots, t_{nn})}{\sum \boldsymbol{a}_{1i} \Delta_i(t_{12}, \ldots, t_{nn})} \right)(\xi)$$

is computed using probabilistic arithmetic.

4 Random linear programming

It is known that for the problem (3)–(4) the optimal solution x^* is achieved at the corner of the set U.

Theorem 3 ([11]). *Let the set U is defined conditions (4), For a point $x = (x_1, ..., x_n) \in U$ is a corner point if and only if there exist numbers $j_1, ...j_r$:*

$$A_{j_1}x_{j_1} + ... + A_{j_r}x_{j_r} = b; x_j = 0, j \neq j_l, l = 1, ..., r,$$

where the columns of the $A_{j_1}, ..., A_{j_r}$ are linearly independent.

For the problem (3)–(4) construct the joint probability density of the vector x^*. For this purpose, we use a method for the solution of deterministic problems of linear programming, for example, the simplex method.

Consider the auxiliary problem. Let A_j, b_j, c_j be sample realizations of the $A_j = A(\omega_j)$, $b_j = b(\omega_j)$, $c_j = c(\omega_j)$, $\omega_j \in \Omega$.

$$\min (c_j, x), \tag{9}$$

s.t.

$$A_j x = b_j, x \geq 0. \tag{10}$$

find a solution x_j^* and the corresponding corner point with numbers$j_1, ...j_r$.

We solve the random system of linear algebraic equations

$$(\boldsymbol{A}_{j_1}...\boldsymbol{A}_{j_r})\boldsymbol{x} = \boldsymbol{b}.$$

If the supports of input parameters are small enough, then due to continuity of $\boldsymbol{x}_j^*$ coincides with $\boldsymbol{x}^*$. In the case of arbitrary supports of input parameters the search procedure for $A(\omega_j), b(\omega_j), c(\omega_j)$, should be repeated, using the Monte Carlo method. In the case that different solutions $\boldsymbol{x}_t^*$ are obtained, they can be compared calculating the probabilistic extension $\boldsymbol{f}_j = (\boldsymbol{c}, \boldsymbol{x}_j^*)$. Notice, that

$$\boldsymbol{x}^* = (\boldsymbol{A}_j)^{-1}\boldsymbol{b}$$

and the expression

$$\boldsymbol{f}_j = \boldsymbol{c}^T(\boldsymbol{A}_j)^{-1}\boldsymbol{b}$$

we can estimate using probabilistic extensions.

Example 1. *As a numerical example, consider the following problem*

$$\min (c, x), \tag{11}$$

s.t.

$$\boldsymbol{A}x = \boldsymbol{b}, x \geq 0. \tag{12}$$

and where $\boldsymbol{A} = (\boldsymbol{a}_{ij})$ is a random matrix, each its element is a random variable with symmetric triangular distribution and support $[\underline{a}_{ij}, \overline{a}_{ij}]$, similarly, $\boldsymbol{b}$ is random vector whose elements are random variables with symmetric triangular distribution.

Support are defined as follows

$$A = \begin{pmatrix} [1-r, 1+r] & [1-r, 1+r] & [3-r, 3+r] & [1-r, 1+r] \\ [1-r, 1+r] & [-1-r, -1+r] & [1-r, 1+r] & [2-r, 2+r] \end{pmatrix},$$

$$b = \begin{pmatrix} [3-r, 3+r] \\ [1-r, 1+r] \end{pmatrix}, c = (-1, -1, 0, 0).$$

For $r = 0$, which corresponds to the deterministic case, the solution is $x^* = (2, 1, 0, 0)$, the columns of the matrix A_1, A_2 correspond to a corner point.

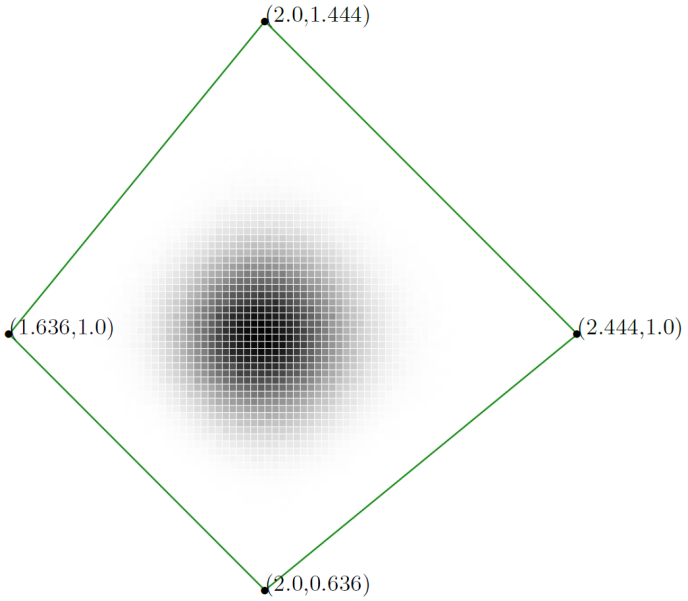


Figure 1: Joint probability density function of the vector (x_1, x_2).

In Fig. 1 the joint probability density function of the vector (x_1, x_2) for $r = 0.1$ with components $x_3 = 0, x_4 = 0$ is shown. The solid line is the boundary of the set of solutions on the (x_1, x_2) plane. The set $\mathcal{X}$ of solutions is the quadrangle with the vertices (2.0,0.636), (2.444,1.0), (2.0,1.444), (1.636,1.0). Value of the probability is represented by shades of gray, as can be seen from the Fig. 1 the probability density is non-uniformly distributed, the highest density is achieved at the center, near the point (2.0, 1.0).

Consider the estimation of the probability density function of the objective function

$$\boldsymbol{f} = c^T (\boldsymbol{A})^{-1} \boldsymbol{b}.$$

Let

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_{11} & \boldsymbol{a}_{12} \\ \boldsymbol{a}_{21} & \boldsymbol{a}_{22} \end{pmatrix}$$

then

$$\boldsymbol{A}^{-1} = \frac{1}{\boldsymbol{\Delta}} \begin{pmatrix} \boldsymbol{a}_{22} & -\boldsymbol{a}_{12} \\ -\boldsymbol{a}_{21} & \boldsymbol{a}_{11} \end{pmatrix},$$

where $\boldsymbol{\Delta} = \boldsymbol{a}_{11}\boldsymbol{a}_{22} - \boldsymbol{a}_{12}\boldsymbol{a}_{21}$. Objective function

$$\boldsymbol{f} = -\frac{1}{\boldsymbol{\Delta}} \left((\boldsymbol{a}_{22} - \boldsymbol{a}_{21})\boldsymbol{b}_1 + (\boldsymbol{a}_{11} - \boldsymbol{a}_{12})\boldsymbol{b}_2 \right)$$

and probabilistic extensions

$$\boldsymbol{f}(\xi) = -\int \boldsymbol{a}_{12}(t_{12}) \dots \boldsymbol{a}_{22}(t_{22}) \frac{1}{\Delta(t_{11}, ..., t_{22})} \left((t_{22} - t_{21})\boldsymbol{b}_1 + (t_{11} - t_{12})\boldsymbol{b}_2\right)(\xi) dt_{12} \dots dt_{22},$$

where $\Delta(t_{11}, ..., t_{22}) = t_{11}t_{22} - t_{12}t_{21}$. The expression

$$\left(\frac{1}{\Delta(t_{11}, ..., t_{22})}((t_{22} - t_{21})\boldsymbol{b}_1 + (t_{11} - t_{12})\boldsymbol{b}_2)\right)(\xi)$$

is computed using probabilistic arithmetic.

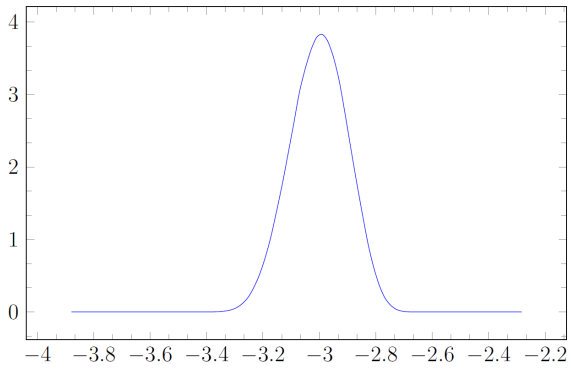


Figure 2: Probability density function of the random objective function (c, x^*)

In Fig. 2 the probabilistic extensions of $\boldsymbol{f}$ is shown.

Conclusion

The considered approach allows us to represent random programming as an effective method for solving linear optimization problems with random input parameters. For this purpose, we used computational probabilistic analysis and procedures for calculating probabilistic extensions. This allowed us to construct a probability density function for the objective function on the set of optimal solutions. The possibility of constructing a probability density function for a set of solutions of systems of linear algebraic equations with random parameters is shown. This approach helps the decision maker to choose the best solutions and allows you to assess the risks.

References

[1] Dobronets B.S., Popova O.A. (2014). Numerical probabilistic analysis under aleatory and epistemic uncertainty. *Reliable Computing*. Vol. **19**, pp. 274–289.

[2] Dobronets, B., Popova, O. (2016). Numerical Probabilistic Approach for Optimization Problems. *In: M. Nehmeier et al. (eds.) Scientific Computing, Computer Arithmetic, and Validated Numerics 2014, LNCS*, Vol.**9553**, pp. 43–53. Springer, Heidelberg DOI: 10.1007/978-3-319-31769-4 4

[3] Dobronets B.S., Popova O.A. (2019). Computational Aspects of Probabilistic Extensions *Tomsk State University Journal of Control and Computer Science*. Vol.**47**, pp. 41–47.

[4] Ewbank J., Foote L. And Kumin H. (1974). A Method For The Solution Of The Distribution Problem Of Stochastic Linear Programming. *SIAM J. APPL. MATH*. Vol. **26**(2), pp. 225–238

[5] Fiedler, M., Nedoma, J., Ramik, J., Rohn, J., Zimmermann K.(2006). *Linear Optimization Problems with Inexact Data*. Springer, Berlin

[6] Kall, P., Mayer, J. (2005). *Stochastic Linear Programming. Models, Theory, and Computation*. Springer, New York

[7] Liu B. (2009). *Theory and Practice of Uncertain Programming (2nd Edition)*. Springer-Verlag, Berlin

[8] Mikhailov G.A., A. Voitishek A.V. (2018). *Statistical modeling. Monte Carlo methods*. Moscow: Urait, (Russian).

[9] Prékopa, A. (1966). On the probability distribution of the optimum of a random linear program. *J. SIAM Control*, Vol. **4**(1), pp. 211–222

[10] Shapiro, A., Dentcheva, D., Ruszczynski, A. (2009). *Lectures on stochastic programming: modeling and theory*. SIAM, Philadelphia

[11] Vasilév, F.P. (1988). *Numerical Methods for Solving Extremal Problems*. Nauka, Moscow (Russian)

Application of artificial intelligence and data compression methods to time series forecasting

Konstantin S. Chirikhin[1] and Boris Ya. Ryabko[2,1]
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *Institute of Computational Technologies SB RAS, Novosibirsk, Russia*
e-mail: chirihin@gmail.com, boris@ryabko.net

Abstract

We consider the problem of time-series forecasting. Nowadays, widespread classical methods can successfully find periodic patterns and simple types of trends. However, in real-life data there may be more complex regularities. For example, time series of economic indicators can have quite complex patterns, because their dynamics can be influenced by participants, who, in turn, can use very complex strategies. In this paper we, first, show how it is possible to combine several methods of forecasting into a single one in such a way that, asymptotically, the accuracy of the obtained method will be equal to the accuracy of the best method among the combined ones. Secondly, we use this approach to combine several methods of forecasting, including "classical" data compression algorithms, a method based on sensing multihead finite automata and a method that uses grammar-based codes. The results of computational experiments are presented.

Keywords: universal coding, time series models, time series forecasting, artificial intelligence.

Introduction

The problem of time series forecasting plays an important role in the study of various economic, physical and social phenomena. Nowadays, there are a huge number of approaches to forecasting, among which we mention neural networks [6], exponential smoothing [5], various models of autoregression and moving averages [2]. But the problem of developing effective forecasting methods is still far from being solved. Classical methods of forecasting can find periodic patterns and basic types of trends well. In [10] it was described how data compression methods can be used to predict time series. Besides periodic patterns, compressors can find "forbidden" combinations (for example, if there is no combination 1111 in a sequence from alphabet $\mathcal{A} = \{0, 1\}$, 0 will be correctly predicted after a subsequence of three 1's). But simple patterns, that are obvious to humans, but cannot be detected by classical methods, still exist. An example of this type of pattern is 01001000100001 ....

In this study we develop a method of time series forecasting which can find some kinds of "difficult" regularities. We use a close connection between problems of data compression and time series forecasting. The approach, described in detail in [11], allows us to efficiently combine various methods of forecasting that present forecast as a probability distribution over possible continuations of the process. It's important

to note that modern data compression algorithms use various heuristics to improve the compression ratio. For example, some compressors search for an approximation of the smallest formal grammar that uniquely describes the sequence for compression. Then, instead of coding the sequence, they encode the found grammar. Thus, we can apply methods with proven efficiency to time series forecasting.

To predict words like 01001000100001... we use the algorithm for sensing multihead finite automaton proposed in [14]. We modify that algorithm in such a way that it can be considered as a data compressor. As a result we obtain a method for forecasting sequences of all mentioned types.

The rest of the paper is organized as follows. In section 1 we discuss how to effectively combine several methods of time series forecasting into a single method. We also briefly discuss the connection between data compression and forecasting. Section 2 describes how context-free grammars can be applied to data compression, and, therefore, to time series forecasting. In section 3 we show how to use a sensing multihead finite automaton together with data compression methods to time series forecasting. In section 4 the results of our computational experiments are presented.

1 An approach to effectively combine methods of forecasting

Suppose that we have a sequence $X = x_1, x_2, \ldots, x_t$, $x_i \in \mathcal{A}$, where $\mathcal{A}$ is a finite set (an alphabet). We consider methods of forecasting which represent forecast of x_{t+1} as a probability distribution over $\mathcal{A}$. Using any such method ϕ we can obtain the probability distribution over $\mathcal{A}^{t+1}$ (the set of all possible sequences of length $t+1$ over alphabet $\mathcal{A}$) as

$$P_\phi(x_1, x_2, \ldots, x_{t+1}) = P_\phi(x_1) \cdot P_\phi(x_2|x_1) \cdot P_\phi(x_3|x_1x_2) \cdot \ldots \cdot P_\phi(x_{t+1}|x_1 \ldots x_t)$$

The estimation of the probability that $a \in \mathcal{A}$ will appear as the next symbol x_{t+1} in a sequence $x_1, x_2, \ldots, x_t$ can be obtained by the following formula:

$$P_\phi(x_{t+1} = a|x_1x_2 \ldots x_t) = \frac{P_\phi(x_1x_2 \ldots x_t a)}{\sum\limits_{b \in \mathcal{A}} P_\phi(x_1x_2 \ldots x_t b)}. \quad (1)$$

If we want to make a forecast for $h > 1$ values ahead, we can obtain a probability distribution over $\mathcal{A}^h$ by using $a, b \in \mathcal{A}^h$ in formula (1).

Using the well-known relation $|\phi(x)| = -\log_2 P_\phi(x)$ between the optimal code length and the probability of a sequence, we can use any data compression algorithm as a predictor. Therefore, in the text below, we will use the terms predictor and compressor interchangeably.

Suppose that we have several algorithms of forecasting $\phi_1, \phi_2, \ldots, \phi_k$ and each of them works well with a specific type of sequence. In this case, we can obtain a single method out of them with almost the same accuracy as the most accurate one for each type of sequence, using the following formula:

$$P_\phi(x_{t+1} = a|x_1x_2\dots x_t) = \frac{\sum\limits_{i=1}^{k} \omega_i 2^{-|\phi_i(x_1x_2\dots x_t a)|}}{\sum\limits_{b\in\mathcal{A}}\sum\limits_{i=1}^{k} \omega_i 2^{-|\phi_i(x_1x_2\dots x_t b)|}}, \tag{2}$$

where the sum of non-negative weight coefficients ω_i is equal to 1. In this study, we used $\omega_i = \frac{1}{k}$.

Thus, the formula (2) allows us to "automatically" select the most appropriate method from the given set. This is due to the fact that the final probability distribution will be mainly influenced by the compressor with the highest compression ratio.

The previously described approach to forecasting can be generalized to forecast real-valued time series. To do so, we can split the set of all reasonably possible values of the time series by m subintervals of equal length, numerate these intervals and then replace the original time series with the series of interval numbers.

Consider the problem of selecting the number of subintervals m. The accuracy and computational complexity of our algorithm strongly depend on this parameter. We consider all partitions into 2^i intervals, $i = 1, 2, \dots, n$, $m = 2^n$. For each partition the forecast is made independently, and then the forecasts are combined with weight coefficients. Denote as x_i an item of the original time series and as y_i^j an interval number corresponding to its value in the partition to 2^j intervals. We can use all partitions with weight coefficients by the formula:

$$P_\phi(y_1^n, y_2^n, \dots, y_t^n) = \frac{\sum\limits_{i=1}^{n} \gamma_i 2^{-|\phi(y_1^i, y_2^i, \dots, y_t^i)| + t(n-i)}}{\sum\limits_{i=1}^{n}\sum\limits_{Z\in\mathcal{A}_i^t} \gamma_i 2^{-|\phi(Z)| + t(n-i)}}, \tag{3}$$

where $\mathcal{A}_i = \{0, 1, \dots, 2^i - 1\}$ is the alphabet of interval numbers and the non-negative weights γ_i sum to 1 (we used $\gamma_i = \frac{1}{n}$). We add $t(n-i)$ bits to the lengths of code words in order to fairly compare them.

2 Forecasting using grammar-based compressors

The concept of grammar-based code was firstly proposed in [7]. The main idea of this approach is to represent a sequence for compression X as a formal grammar G using which the original sequence could be uniquely derived: $L(G) = \{X\}$, where $L(G)$ is the set of all sequences that can be derived from G (the language defined by G). To the purpose of data compression it is desirable to find as compact G as possible. In such a general formulation, the problem is reduced to the calculation of the word's Kolmogorov complexity, which, as known, is not computable. Therefore, the class of considered grammars is usually limited to context-free grammars (CFG). The problem of searching the smallest CFG representing given sequence is known

as the smallest grammar problem (SGP). In paper [3] it was proved that SGP is NP-complete, so we have to use approximation algorithms. Once G is found, instead of compressing the original sequence, we can compress G using arithmetic coding or some special algorithm, designed to compress formal grammars. We can illustrate the basic idea by a simple example. Suppose we want to compress the sequence 10011000100010011. We can represent it with the following context-free grammar:

$$S \to X_1X_2X_2X_11$$
$$X_1 \to 1001$$
$$X_2 \to 1000$$

A variety of grammar-based compressors have been developed, among which we mention Sequitur [9] and lca [12], which are both based on context-free grammars.

To the purpose of forecasting we can use any grammar-based compressor in formula (2).

3 Forecasting using sensing multi-DFAs

In this section, we consider how to predict words like 010010001 . . . using a compression-based approach. Words of this type belong to the class of multilinear words, defined in [13]. An infinite word is multilinear, if it can be written in the form:

$$q\prod_{n\geq 0} r_1^{a_1n+b_1}r_2^{a_2n+b_2}\dots r_m^{a_mn+b_m},$$

where $\prod$ denotes concatenation, q is a some finite word, m is a positive integer, and for each $1 \leq i \leq m$, r_i is a nonempty word, a_i and b_i are non-negative integers such that $a_i + b_i > 0$.

To proceed further, we need the following definitions. Multihead deterministic finite automata (multi-DFA) predictor is a tuple of the form $M = (Q, \mathcal{A}, k, T, \triangleright, q_S)$, where Q is the finite set of states, $\mathcal{A}$ is the input alphabet, $k \geq 1$ is the number of input heads, $\triangleright$ is the start-of-input marker, a_S is the initial state, and T is a transition function of the form:

$$T : [Q \times (\mathcal{A} \cup \{\triangleright\})^k] \to [Q \times \{\text{stay}, \text{right}\}^k \times \mathcal{A}].$$

A sensing multi-DFA is an extension of multi-DFA, in which its transition function takes an additional argument indicating, for each pair of heads, whether those two heads are on the same input position.

In [14], an algorithm of forecasting multilinear words was proposed. More precisely, it was proved that there is a sensing multi-DFA which masters every multilinear word over alphabet $\mathcal{A}$ and an algorithm which it should implement was provided. In this study we use that algorithm to obtain a probability distribution over $\mathcal{A}^n$. As a result, we can use the automaton as an ordinary data compressor in formulas (1-3).

In this paper we don't repeat the description of the original algorithm from [14] and describe only our modifications. The basic idea is as follows. Suppose that a word $x_1x_2 \dots x_n$ is given, and we have processed a prefix of this word $x_1x_2 \dots x_t$,

$t < n$. The original algorithm attempts to exactly predict the value of the next term x_{t+1}. We modify this algorithm so that it gives a probability distribution of the next symbol using the already processed part of the series $x_1, x_2, \ldots, x_t$. The original algorithm can be decomposed into two stages:

1. *Correction* — at this stage, the automaton tries to adjust its heads and in most situations cannot give a reasonable prediction. In such situations, any value can be used as the prediction in the original algorithm;

2. *Matching* — the automaton works under assumption that all heads are positioned correctly. The automaton gives a certain letter as a prediction.

In the first case, we obtain the probability distribution of the next symbol using the so-called Krichevsky predictor [8]:

$$P_A(x_{t+1} = a | x_1 \ldots x_t) = \frac{\nu_{x_1 \ldots x_t}(a) + 1/2}{t + |\mathcal{A}|/2},$$

where $\nu_{x_1 \ldots x_t}(a)$ is the number of occurrences of letter a in the word $x_1 \ldots x_t$.

In the second case, we want to give more weight to the prediction of the automaton. In addition, the more correct predictions were made by the automaton, the more weight should be given to its prediction. Therefore, we used the following heuristic formula:

$$P_A(x_{t+1} = a | x_1 \ldots x_t) = \frac{C + 1/2}{C + |\mathcal{A}|/2},$$

where C is the length of the continuous series of correct predictions by the automaton up to the current symbol (if $C = 0$, we have a uniform distribution). We assign the same probability to the remaining letters.

4 Computational experiments

In this section, we present the results of an experimental study of our algorithm. We used both well-known file compression programs and our implementation of the automaton. The following programs were used:

- *zlib*, version 1.2.11. It can be downloaded by the link *https://www.zlib.net*;
- *ppmd*. We used the implementation of E. Shelwien which can be downloaded by the link *https://github.com/Shelwien/ppmd_sh*;
- *rp*. We used the implementation described in [1];
- *automaton*, our own implementation.

We begin our description with an artificial sequence which is a prefix of the proper multilinear word:

$$12345112233445111222333444 5\ldots = \prod_{i=1}^{\infty} 1^i 2^i 3^i 4^i 5 \tag{4}$$

We call a sequence of 1s, 2s, 3s, 4s, followed by 5, as a block. The results of our computations are presented in table 1 (comb. means that zlib, ppmd, rp and automaton were used as a single method). In table 1, the notation 'a/b' means that we used the prefix $x_1, x_2, \ldots, x_a$ of word (4) as an initial history to make b one-step forecasts (for $x_{a+1}, x_{a+2}, \ldots, x_{a+b}$). The error of a method was computed as the sum of the modules of the b individual errors.

Table 1: The results of forecasting series (4)

History/Forecast	zlib	ppmd	rp	aut.	comb.
230/200 (10 blocks)	128	70	148	113	70
860/200 (20 blocks)	105	52	153	0	52
5150/200 (50 blocks)	104	51	124.0	0	51
45450/200 (150 blocks)	97	32	171	0	22
80600/200 (200 blocks)	79	12	201	0	0

As can be seen from table 1, the automaton is the best forecasting method for that sequence. In addition, with increasing sequence length, the accuracy of the combined method becomes the same as the accuracy of the automaton.

Let us consider another example. To show that our method can be used to forecast real-world data, we performed computations for the sunspot number time series. The Royal Observatory of Belgium (WDC-SILSO, Royal Observatory of Belgium, Brussels, *http://www.sidc.be/SILSO/*) publishes on its website monthly mean sunspot number data along with the forecasts. We performed our computations as follows. For each month from August 2015 to September 2018, we made forecasts for 18 months ahead. Then from the file 18 month later, we took the observed values and computed the mean absolute error (MAE):

$$\mathrm{MAE} = \frac{1}{h}\sum_{i=1}^{h} |\hat{x}_i - x_i|,$$

where $\hat{x}_i$ is the predicted value, and x_i is the observed value.

We used the following time series preprocessing techniques:

1. To remove the seasonal component from the series we used seasonal trend decomposition (STL) [4]. The frequency of the seasonal component was equal to 11 years or 132 month (because the duration of the solar cycle is typically 10-11 years);

2. We used smoothing function: $x_t^* = (2x_t + x_{t-1} + x_{t-2})/4$;

Table 2: The results of the sunspots number forecasting

	MAE for forecasting horizon										Average		
	1	2	3	4	5	6	8	12	15	18	1-4	1-8	1-18
SILSO	10.5	11.4	11.8	12.7	13.9	14.9	16.5	15.9	17.1	20.3	11.6	13.5	15.5
zlib	10.5	18.2	28.7	16.9	10.6	10.8	28.9	12.5	19.2	16.4	18.6	17.4	19.2
ppmd	10.7	12.9	12.3	10.9	15.3	12.0	25.0	12.9	9.4	16.7	11.7	13.9	15.8
rp	17.4	34.7	57.8	31.0	34.2	29.9	65.5	46.7	65.7	36.1	35.2	38.8	49.1
aut.	26.6	53.2	78.9	52.2	54.9	8.9	114.3	10.2	185.1	10.1	52.7	60.0	92.9
comb.	10.7	12.9	12.3	10.9	15.3	12.0	25.0	12.9	9.4	16.7	11.7	13.9	15.8

3. We split the time series to the 6 time series (so instead of making one forecast for 18 values ahead, we made 6 forecasts for 3 values);

4. We took a first difference, i.e. instead of forecasting a series $x_1, x_2, \ldots, x_t$ we considered the series $x_2 - x_1, x_3 - x_2, \ldots, x_t - x_{t-1}$;

5. We considered partitions to 2, 4, 8 and 16 intervals.

Since SILSO forecasts are made for a smoothed time series, they are 6 months late due to the smoothing method. Therefore, we used their forecasts from 7 to 25 values ahead. The results of computations are presented in table 2.

As we can see from table 2, our method has comparable accuracy with the method of SILSO.

Conclusions

In this paper was presented an algorithm of time series forecasting, which can capture some kinds of "difficult" regularities. By performing calculations on both artificial and real-world data, we showed that the algorithm can find such regularities if they are present in the data. Otherwise, our modifications do not degrade the accuracy of classical data compression methods.

Acknowledgments

This work was supported by Russian Foundation for Basic Research (grant 19-47-540001).

References

[1] Bille P., Gørtz I.L, Prezza N. (2017). Space-efficient re-pair compression. *In Data Compression Conference (DCC). IEEE.* pp. 171-180.

[2] Box G.E., Jenkins G.M., Reinsel G.C., Ljung G.M. (2015) *Time series analysis: forecasting and control*. John Wiley & Sons.

[3] Charikar M., Lehman E., Liu D., Panigrahy R., Prabhakaran M., Sahai A., Shelat A. (2005) The smallest grammar problem. *IEEE Transactions on Information Theory*. Vol. **51**, pp. 2554-2576.

[4] Cleveland R.B, Cleveland W.S., McRae J.E., Terpenning I. (1990) STL: A seasonal-trend decomposition. *Journal of official statistics*. Vol. **6** pp. 3-73.

[5] Hyndman R., Koehler A.B., Ord J.K., Snyder R.D. (2008) *Forecasting with exponential smoothing: the state space approach*. Springer Science & Business Media.

[6] Kaastra I., Boyd M. (1996) Designing a neural network for forecasting financial and economic time series. *Neurocomputing*. Vol. **10** pp. 215-236.

[7] Kieffer J.C., Yang E.H. (2000) Grammar-based codes: a new class of universal lossless source codes. *IEEE Transactions on Information Theory*. Vol. **46** pp. 737-754.

[8] Krichevsky R. (1968) A relation between the plausibility of information about a source and encoding redundancy. *Problems Inform. Transmission*. Vol. **4** pp. 48-57.

[9] Nevill-Manning C.G., Witten I.H. (1997) Identifying hierarchical structure in sequences: A linear-time algorithm. *Journal of Artificial Intelligence Research*. Vol. **7** pp. 67-82.

[10] Ryabko B.Y. (1988) Prediction of random sequences and universal coding. *Problems of information transmission*. Vol. **24** pp. 87-96.

[11] Ryabko B., Astola J., Malyutov M. (2016) *Compression-based methods of statistical analysis and prediction of time series*. Springer International Publishing.

[12] Sakamoto H., Kida T., Shimozono S. (2004) A space-saving linear-time algorithm for grammar-based compression. *In International Symposium on String Processing and Information Retrieval* pp. 218-229.

[13] Smith T. (2013) On infinite words determined by stack automata. *In IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2013)*

[14] Tim S. (2018) Prediction of infinite words with automata. *Theor. Comp. Sys.* Vol. **62**, pp. 653-681.

Approaches to customers lifetime value prediction

Natalia Galanova
Skyeng, Novosibirsk, Russia
e-mail: natalia.galanova@gmail.com

Abstract

This paper considers various approaches to the problem of customers' lifetime value predictions. We discuss several approaches that can be used in the non-contactual settings - Pareto/NBD model, survival analysis approaches and classical regression techniques. Also we discuss the advantages and disadvantages for each group of methods and the challenges that we face in case of using survival analysis for customers' lifetime prediction.

Keywords: lifetime, lifetime value, LTV, CLV, Pareto/NBD model, Cox model, survival analysis, censored data.

Introduction

In marketing, customer lifetime value (or LTV) is a prediction of the net profit attributed to the entire relationship with a customer. The reasons to predict customers' lifetime values may be various, these are the most popular:

- understand difference between user's acquisition cost and his future profit;
- find clusters among users, based on their lifetime;
- focus attention and support on users with long-term LTV forecasts;
- find features and characteristics that affect (or at least correlate with) LTV forecast value.

Skyeng is the online English language school and it is the biggest online school in Russia - we have about 8000 teachers from all over the world and more than 73000 of students, mostly from Russia and eastern Europe.

In our internal processes prediction of student's lifetime and lifetime value is a very important goal - because we have different costs for customer's acquisiton and different costs for customer's support. And if we have some amount of users, which are suggested to churn in the near future - it would be better for us to understand the total profit for each student so we could be able to prioritize our efforts to keep these customers.

Also it should be noted that in this paper we speak about lifetime and lifetime value predictions for companies with non-contractual settings - when the is no contract between customer and company and customer can make a purchase at any time.

In case of Skyeng, we have students and we do not control their intensity of studying - so we have non-contractual settings and consider students' lessons as their random purchases.

There are several approaches to predict lifetime value:

1. we can consider customer's lifetime value (or customer's lifetime) as a random variable and use various regression techniques for prediction (starting from simple linear regression and ending up with gradient boosting on decision trees models or even neural networks);

2. we can use specific models for lifetime and lifetime value prediction, such as RFM-models (recency-frequency-monetary models), which will be discussed below;

3. we can use survival models for predicting customers' lifetime (in terms of number of transactions) and then recalculate it to the lifetime value, based on the price distribution for each customer.

Approach with forecasting lifetime value by lifetime prediction is only valid if we are sure that price distribution for each customer is normal (or at least close to normal) and has small variance. In our case we can use such an approach, because the difference in costs per lesson is small and also we usually have some prior information about the reasons that affect the prices for particular student:

- what kind of teacher student has - native or not (because native is more expensive)?

- what was the amount of the last payment (higher payments give you bonuses and lower prices for each lesson)?

- what were the amounts of all previous payments (to predict future ones)?

Moreover - the approach with prediction lifetime value by lifetime is more stable, especially if we are going to change prices for our products in future or even change the strategy of price formation. That is why we will consider only this way of LTV prediction.

Pareto/NBD model

The Pareto/NBD model is perhaps the most well-known and frequently applied probabilistic model in the non-contractual context [3].

The Pareto/NBD model makes the following assumptions regarding the customer population:

- Purchase count follows a Poisson distribution with rate λ. In other words, the timing of these purchases is somewhat random, but the rate (in counts/unit time) is constant. In turn, this implies that the inter-purchase time at the customer level should follow an exponential distribution;

- Lifetime distribution follows an exponential distribution with slope μ. The expectation value of such distribution is $\frac{1}{\mu}$ and corresponds to the lifetime of the user;

- The latent parameters λ and μ are constrained by two prior gamma distributions representing our belief of how these latent parameters are distributed among the population of customers. These two gamma distributions have parameters (r, α) for the purchase count and (s, β) for the lifetime. The goal is to find these four parameters. From these, all actionable metrics can be derived.

The Pareto/NBD model focuses only on modeling lifetime and purchase count, but LTV involves not only lifetime but also the cost of every purchase. So there is monetary value extension to the Pareto/NBD model - so called Gamma-Gamma model [2].

Survival models

Another approach to LTV prediction is to use survival analysis techniques, e.g. famous Cox model [1]. In that case we should consider the number of customer's purchases as positive random variable - our time-to-event observation.

But there are some differences between customers LTV data in comparing with classical time-to-event data. When we speak about LTV in marketing or e-commerce data usually consists of user's transactions in time and each transaction has its own cost (in some cases costs may be the same). The count of these transactions per user we can consider as user's lifetime and the sum of these transactions we can consider as user's lifetime-value. So we have some important differences in data, that can be explained with the Figure 1, which shows a few purchasing trajectories for different types of customers. Time goes from left to right, the vertical dashed line represents the present time, and each small, vertical line represents a purchase/lesson made by a customer:

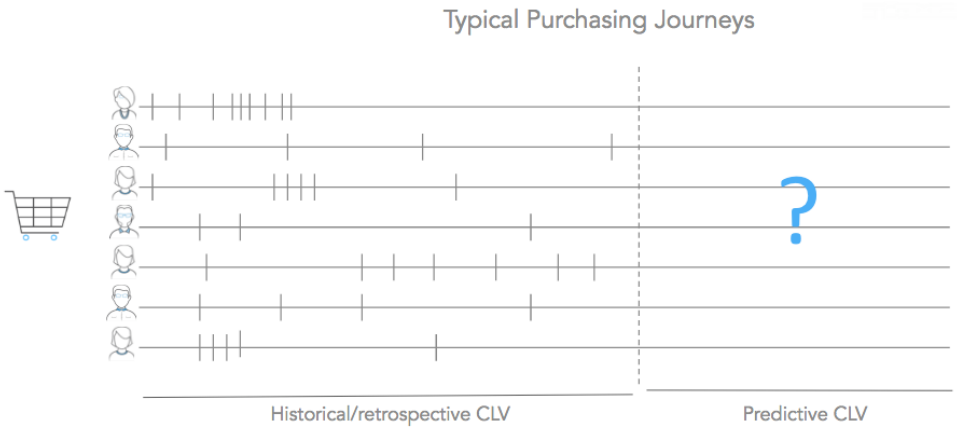


Figure 1: Customers' puchases

The main thing is that we don't know exactly if the 'failure' has occurred. As failure-event we usually consider the fact that user has churned (stopped any future

relationships with the company). All that we know is the time from the last transaction/lesson and based on this information we need to decide whether this user has churned or not. Usually companies have some fixed period based on historical data or any past experience and after that period customer is considered as churned user. But the obvious disadvantage of this approach is illustrated at the Figure 1 - customers can have very different frequency of purchases and one fixed value for everybody can be a bad solution. So in our work we decided to use different technique, which is based on the previous intensity of purchases for each customer. For instance, we can use Pareto/NBD model, which was discussed above, for predicting probability of churn or 'failure' event.

Another problem is that we have a lot of censored observations - our active students. And in case of prediction these censored data are more important than our churned users, which are complete observations - because newer students have the most relevant features and face the most relevant business processes and for churned users all the covariates can be irrelevant.

So we need to achieve a balance here - we can't neglect censored observations, because of its importance and on the other hand, we can't use all of the censored data in models, because it will lead to poor parameters' estimation.

Comparison of the approaches

Table 1: Comparison of the approaches

	Pareto/NBD	Survival analysis	Regression techniques
LT/LTV	can be used both for lifetime and lifetime value predictions	supported by the nature of the models	unsupported - we can't add any covariates to these models
Censored data	can be used only for lifetime prediction and then we need to recalculate to LTV	supported, but we need to find a balance between censored and completed data	supported - we can add any covariates, even time-varying to these models
Covariates	can be used both for lifetime and lifetime value predictions	unsupported (so for prediction we have only churned and possibly irrelevant users)	supported - but mostly we can add only time-fixed covariates to these models

Let us consider the main advantages and disadvantages of all discussed methods. The Table 1 presents features of the each approach for three points of view:

- can we use it for predicting lifetime value or only lifetime?
- can we use censored data? because as it was mentioned above - censored data are the most relevant for us;
- can we use covariates? because besides student's intensity of lessons we have a lot of features like age, gender, country, level of english and so on.

Conclusions

We discussed why the companies need to forecast customers' lifetime values and the problems that they face in making these predictions. We also discussed three main groups of methods for solving that problem, its advantages and disadvantages. And as a conclusion we can say that only the methods of survival analysis can provide the most correct results - as a combination of using covariates and censored (the most relevant) data.

References

[1] Cox, D. R.; Oakes, D. (1984). *Analysis of Survival Data*. Chapman and Hall, London – New York

[2] Fader P. S., Hardie B. G. S., Lee K. L. (2005). "Counting your customers" the easy way: An alternative to the Pareto/NBD model *Marketing Science*. Vol. **24**, pp. 275-284.

[3] Schmittlein D. C., Morrison D. G., Colombo R. (1987). Counting your customers: Who are they and what will they do next? *Management Science*. Vol. **33**, pp. 1-24.

About the task of leveling the “false” operations of the heat load regulator

Nadezhda V. Kononova[1], Denis A. Zhalnin[2], Olesya V. Chubarova [2]
[1] *Siberian Federal university, Krasnoyarsk, Russia*
[2] *Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russia*
e-mail: koplyarovanv@mail.ru, region-ingener@yandex.ru, Kuznetcova_o@mail.ru

Abstract

The effect of a heat load regulator «false» operation is inherent in systems with cross connections (boilers — main steam main — turbine, fast-response PRDS - fast-response pressure-reducing and desuperheating station). The boiler heat load regulator (HLR) is triggered «falsely» to a significant external disturbance from the steam line. The work is devoted to the problem of a boilers HLR response correction solving for the regulator «false» operation leveling for a thermal power plant (TPP) with cross links (or the combined heat and power plant (CHP)). The solution is based on the analysis of the data from measurement sensors in real time.

Keywords: coal thermal power station, combined heat and power plant, steam line, heat load regulator, «false» operations, algorithm of leveling, correction circuit.

Introduction

Thermal power plant with cross connections is a system of boilers and turbines connected by a single heat line. Boilers produce superheated steam, which is collected in the main steam line. Next, steam enters for heating system water and the turbine to produce electricity. In such a unified system, control of each boiler separately is sensitive to external disturbances and to changes in the station load.

To maintain a given level of steam consumption, an automatic control loop is implemented - a heat load regulator (HLR). The difference between the current value of the steam flow and the target flow is supplied to the controller input. In some solutions, a pressure change in the steam collecting chamber is additionally added through the adder.

The scheme and experience of implementing this solution is described in detail in [1]. The heat load regulator of the boiler unit is triggered “falsely” with a significant external disturbance from the steam main. Such a disturbance may be caused by a change in the load of the turbines, that is, the execution of the dispatch schedule and is the normal operation of the station. The situation is accompanied by a change in the load on the boiler and turbines. In the case when a part of the boiler remain with its previous load, then their HLR will "falsely" work out, which will lead to an increase in the total time for reaching the steady state. Since the system of boilers, pipelines and

turbines is a system of interconnected vessels, where one affects all, “false” operations introduce additional disturbances to the natural pressure fluctuations in the pipeline and make it difficult to stabilize the entire system. The physical essence of the processes occurring at the time, and the reasons for the appearance of such values of consumption on the sensors were described in detail by A.S. Klyuev in «Installation of Systems for Automatic Regulation of Drum Steam Boilers» [2].

The «false» operation of the regulator effect results are:

- increased risk of pressure fluctuations (higher than desired);
- with regular “false” border operations (feasible area) are observed in a wider range than recommended;
- increase the total time to enter the mode;
- increased equipment wear.

Today the solution of "false" operation problem is to change the HLR task and the raw coal feeder speed by operator in manual mode.

1 Problem formulation

A system of 4 boilers of CHP included in a common steam main is considered as a research object.

Let there is a boiler unit (object of study), which is part of an interconnected system (group of boiler units, the main steam line and turbines).

The boilers structures are the same. Each boiler has a set of regulators (the level in the boiler drum, vacuum, etc.), which have the task of maintaining or controlling certain variables. We should note, that boilers has the heat load regulator (PI controller), that controls its steam capacity.

Let a disturbance occurs in the system (change in turbine load). In the same time the task for the studied object (boiler) does not change.

In this case one of the main control system disadvantages is the false operation of the heat load controller of boiler in the base mode as the reaction to perturbations from the main steam line side.

It is required to form a control action on the object in such a way as to minimize the effect of “false” triggers of the heat load regulator of the considered boiler unit.

2 The false HLR operation leveling algorithm

To solve the problem, an algorithm for leveling “false” heat load regulator operations in automatic mode is proposed.

To do this, we introduce into the control circuit an algorithm (named control unit) that analyzes the general situation of the station load and levels the “false” HLR responses if it is nessasary.

For boilers with constant load it leaves the current number of the raw coal feeder revolutions, that ensuring their load in the steady state. The state stabilization will be the task of the boiler units, which introduce disturbances, that is, change the load.

The integration scheme of the data processing algorithm requires the complete preservation of the existing station control scheme, so that the basic control loop is always included in the work.

The solution scheme is presented in the figure (1).

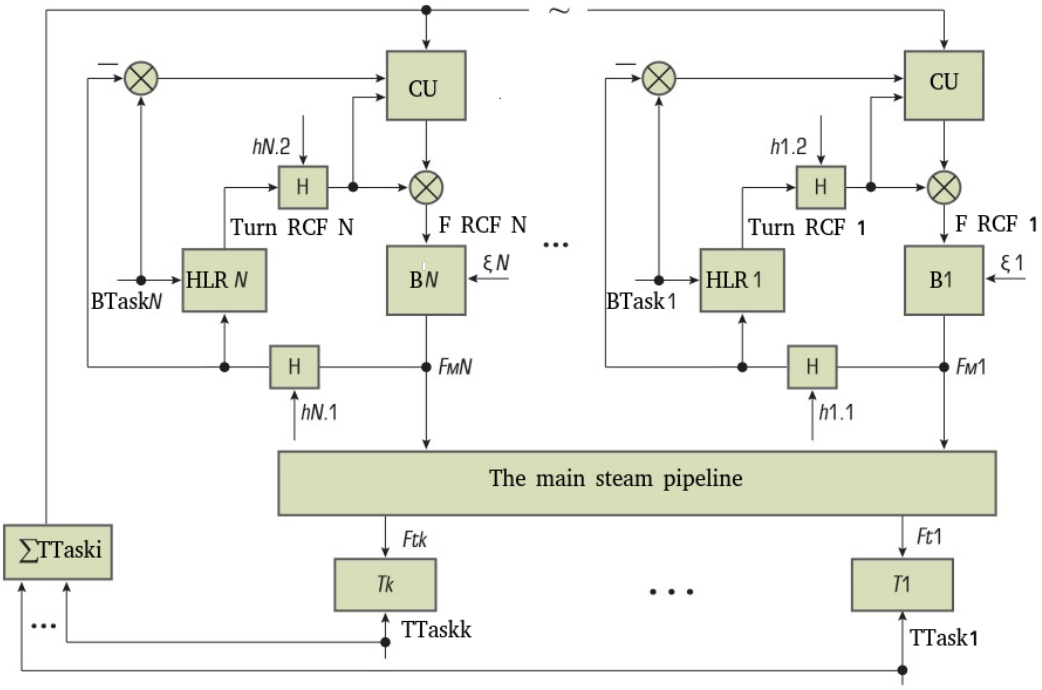


Figure 1: The scheme of the rpm number of raw coal feeder correction (by the proposed control unit) with the HLR enabled

The following notation are used in the diagram: $B(1, \ldots, N)$, $T(1, \ldots, k)$ – boilers and turbines; control unit (CU) – the proposed algorithm for calculating the raw coal feeder (RCF) rpm number in the boiler , taking into account the load change throughout the station; HLR$(1, \ldots, N)$ – classical proportional-integral (PI) regulator of heat load (for each boiler); BTask$(1, \ldots, N)$ – steam consumption task in each boiler; TTask$(1, \ldots, k)$ – steam consumption task for each turbine; ΣTTaski - total task for turbines; Ft$(1,\ldots, k)$ – measuring steam consumption turbines; Turn RCF $(1, \ldots, N)$ – rpm number of raw coal feeder from boiler HLR; F RCF $(1, \ldots, N)$ – corrected raw-coal feeder turnover values; $\xi(1, \ldots, N)$ – uncontrolled disturbances (changes in fuel quality, air regime, thermal properties of feed water, the operation of the superheated steam temperature controller, etc.); hN.1, (N – boiler number) – steam flow fluctuations, measurement accuracy; hN.2, (N – boiler number) – the conversion of the output signal in the RCF rotation frequency.

In the base control scheme PI regulators control the values of the steam flow by changing the input influences (rpm number of raw coal feeder).

The proposed control unit (CU) monitors the current situation according to the process data and, if necessary, includes an algorithm for leveling “false” operations.

The leveling algorithm is activated in the boiler where an external disturbance by other boilers is detected when the station load changes. It works only for boilers where HLR task remains unchanged.

The leveling "false" regulator operations [4] algorithm consists of the following steps:

1. Fix a change of the turbine task;

2. Trace changes of the task to boiler units;

3. If the boiler has not received a change in steam flow task, then switch off it HLR

4. The control unit supplies the input of the boiler with the value of the RCF revolutions from the previous step (before the turbine load changes).

5. After the end of the transient processes in boilers with a modified load, the control unit turns on the RCF in the boiler without changing the load with zero value of the integrating component.

The main feature of the solution is the use of current information, that is, the latest measurements from sensors throughout the station to correct the reaction of each boiler, which is expected to have a “false” operation of the regulator. In this regard, it is important to ensure the flow of raw data from sensors with reliable values into the algorithm, that requires the information-measuring system and instrumentation stability.

In developing and testing the HLR operation correction algorithm, raw data archives were used. Data was written to the file in increments of 30 seconds.

The initial archive was the unloading of process control parameters, in *.mbd, MS Access format. Data acquisition time: from june 4, 2017 to december 14, 2017. The data processing period is taken from June 5 to June 27, 2017.

List of the thermal power plant variables [3]:

- the coal consumption in boiler: 0-6000 a rpm number of raw coal feeder.

- the steam pressure in the drum is in range 120-150 kgf/cm^2;

- the steam flow from boiler: 0-600 t/h

- the steam pressure in the manifold (main line): 120-150 kgf/cm^2;

- the steam flow in the manifold (main line): 0-1700 t/h ;

- the heat load regulator task (flow task) in boiler: 0-600 t/h. (the max value of flow task for different boilers is 350 or 500 t/h)

- the total number of fuel turns for monitoring the task change process: 0-24000 rpm.

Some computational studies of the imitation object are carried out in the paper.

3 Test experiments on the object model

To test the proposed algorithm of the HLR false operations correcting, an experiment was conducted. On the ABB (Asea Brown Boveri Ltd.) platform a prototype of the station with 4 boilers combined into a single steam main was created (see in figure (2)).

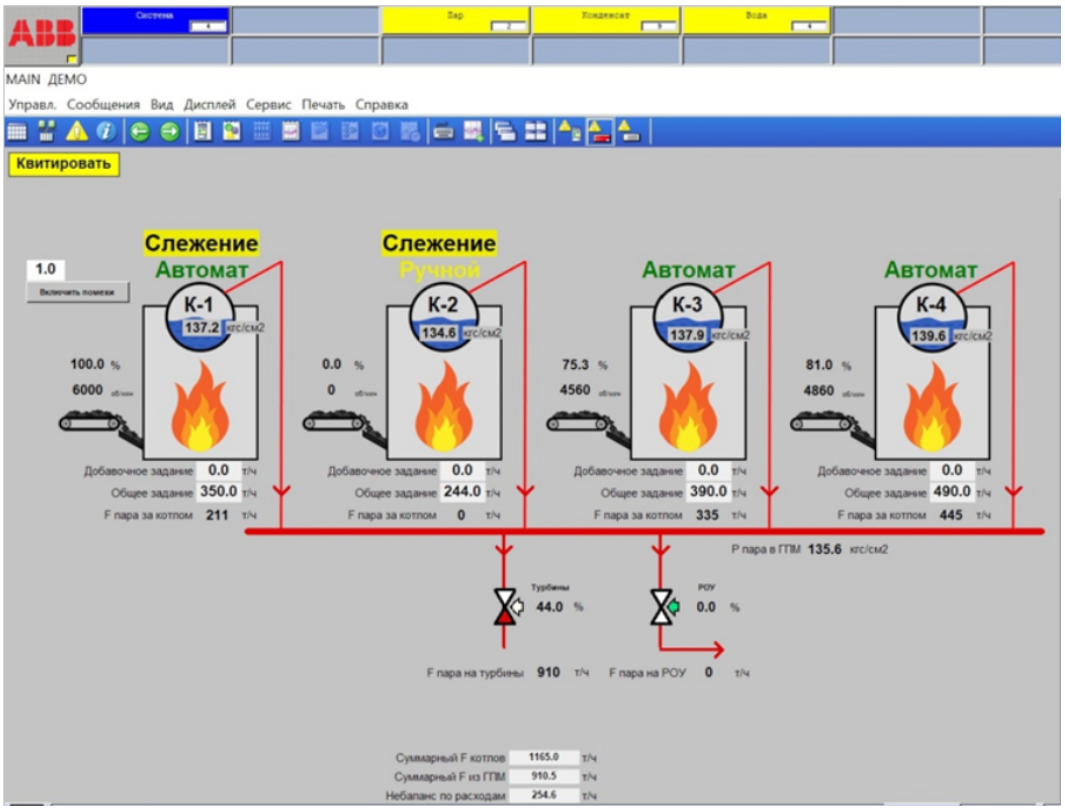


Figure 2: Imitation model of the station with 4 boilers

The simulation model is based on the following dependencies derived from real data:

1. Main steam line pressure is calculate as:

$$P_{MSL} = ((F1 + F2 + F3 + F4 - Ft - F_{PRDS})40e^{-st}\frac{1}{(Ts+1)}\frac{1}{5Ts}) + 125. \quad (1)$$

where P_{MSL} - steam pressure in the main steam line, $F1$ – steam flow of boiler 1; $F2$ – steam flow of boiler 2; $F3$ – steam flow of boiler 3; $F4$ – steam flow of boiler 4, Ft - turbine steam flow, F_{PRDS} – pressure-reducing station flow.

2. To calculate the pressure of each boiler separately the following formulas are used:

$$P_{drm1} = (3.7V1^2 0.0000001\frac{1}{(Ts+1)(Ts+1)}) + \frac{P_{MSL}}{(Ts+1)}. \quad (2)$$

where P_{drm1} – drum pressure 1, $V1$ - a rpm number of the heat load regulator 1;

$$P_{drm2} = (3.8V2^2 0.0000001\frac{1}{(Ts+1)(Ts+1)}) + \frac{P_{MSL}}{(Ts+1)}. \quad (3)$$

where P_{drm2} – drum pressure 2, $V2$ - a rpm number of the heat load regulator 2;

$$P_{drm3} = (3.9V3^2 0.0000001 \frac{1}{(Ts+1)(Ts+1)}) + \frac{P_{MSL}}{(Ts+1)}. \quad (4)$$

where P_{drm3} – drum pressure 3, $V3$ - a rpm number of the heat load regulator 3;

$$P_{drm4} = (4V4^2 0.0000001 \frac{1}{(Ts+1)(Ts+1)}) + \frac{P_{MSL}}{(Ts+1)}. \quad (5)$$

where P_{drm4} – drum pressure 4, $V4$ - a rpm number of the heat load regulator 4.

3. To calculate the steam flow of each boiler separately the following formulas are used:

$$F_1 = 170\sqrt{P_{drm1} - P_{MSL}}, \quad (6)$$

$$F_2 = 170\sqrt{P_{drm2} - P_{MSL}}, \quad (7)$$

$$F_3 = 223\sqrt{P_{drm3} - P_{MSL}}, \quad (8)$$

$$F_4 = 223\sqrt{P_{drm4} - P_{MSL}}. \quad (9)$$

4. To calculate the steam flow of turbines and pressure-reducing station flow the following formulas are used:

$$F_T = 177.7K_{VFC.Turb}\sqrt{P_{MSL}}. \quad (10)$$

$$F_{PRDS} = 9K_{VFC.PRDS}\sqrt{P_{MSL}}. \quad (11)$$

where $K_{VFC.Turb}$ - valve flow coefficient of turbine, $K_{VFC.PRDS}$ - valve flow coefficient of pressure-reducing station.

5. The proportional-integral regulator for HLR imitation is described as:

$$W(s) = \frac{Kp(Ts+1)}{Ts}. \quad (12)$$

where Kp, Ts - parameters of PI regulator.

It should be noted that the efficiency of the control with the use of PI regulator depends on when and how accurately it is set. Since over time some characteristics of the object may change, which will require the setting of the PI controller parameters.

Next, we show the result of the experiment with the leveling Control unit turned on and without it on the imitation model (figures (3), (4)).

We compare the results of steam flow control in the main line with the use of the PI-regulator and the proposed controller that corrects the reaction of the boiler heat load controller. It is required to bring the system to the setpoint state when changing the target for steam flow in the pipeline, by controlling the heat load controller of the one boiler. The remaining boilers of the CHP plant are in the base mode (and their regulators react to the pipeline pressure changes).

Figure (3) shows the difference in the rpm number of raw coal feeder (RCF) with the control unit turned on (algorithm for leveling "false" operations) and without it.

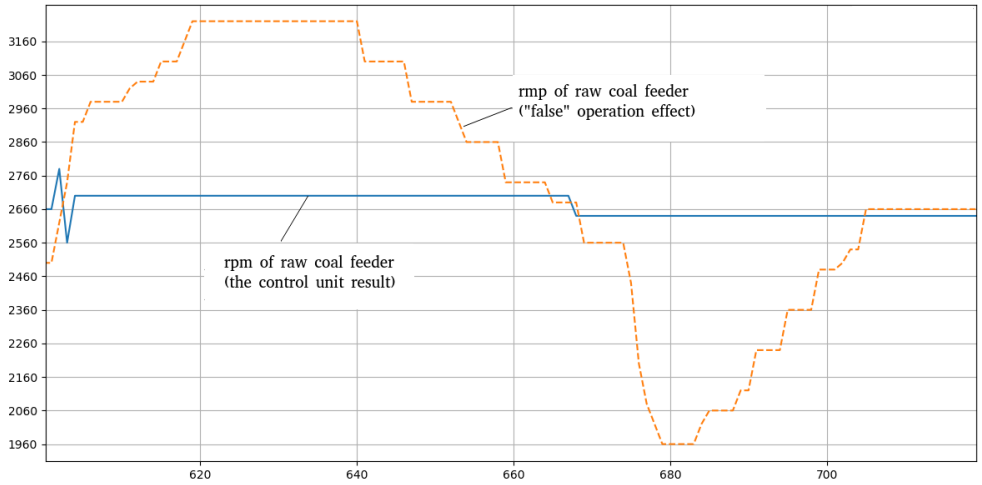


Figure 3: An example of the RCF speed change with control unit operation and without it

Figure (4) illustrates the difference in the operation of the boiler with the leveling algorithm turned on and without it.

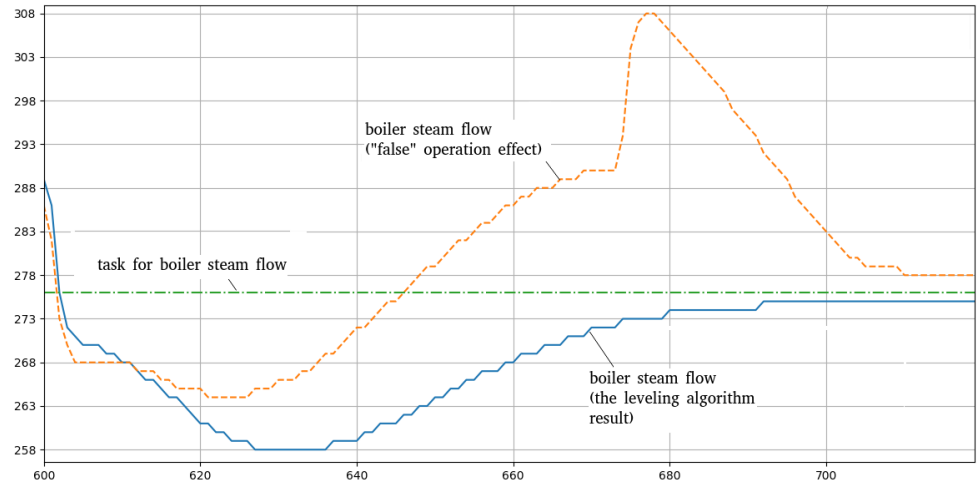


Figure 4: The Boiler steam flow with and without control unit operation

As we can seen from the figure (4), variations in steam flow at the output of the boiler with the leveling algorithm turned on are much lower. Thus, the use of the leveling algorithm allows us to avoid additional pressure fluctuations introduced into the system under consideration in the presence of "false" regulator operations, as well as to reduce the time for the entire system to achieve the specified performance.

Conclusions

The introduction of the proposed algorithm in the control system at the plant will allows correcting the false reaction of the heat load regulator of boilers in order to achieve the setpoint.

The proposed algorithm allows to avoid the effect of "false" operation, calculates the required value of the control action. It blocks the control signal from the heat load regulator by transferring the calculated values, then returns control to it, zeroing the integrating component. As a result of all the above actions, the control unit helps prevent unwanted pressure fluctuations in the main steam line.

References

[1] Zhalnin D.A., Shorokhov V.A., Evdokimov A.N., Bubnovskii O.A., Churinov A.V.(2010) Experience with the introduction of a pressure regulator system in the main steam line at the Krasnoyarsk TÉTs-2 plant. [in Russian]*Power Technology and Engineering*, Vol. **11**, pp. 18-25.

[2] Klyuev A.S. (1985) Installation of Systems for Automatic Regulation of Drum Steam Boilers [in Russian], // Klyuev A.S., Lebedev A.T., Novikov S.I. Moscow, Énergoatomizdat, 1985, 280 p.

[3] Koplyarova. N. (2017) About the Control of a Group of Objects on the Example of Steam Pressure in the CHP Main Line // N. Koplyarova, A. Chubarov, N.Sergeeva *AMSA'2017*, Krasnoyarsk, Russia, 18-22 September, 2017: Proceedings of the international workshop. – Novosibirsk: NSTU publisher, 2017, pp. 96-103.

[4] Sergeeva N.A.(2018) Application of current data for operational control [in Russian] // N.A. Sergeeva, O.V. Chubarova, *Modern methods and algorithms of automation in power engineering*, 2018, №1 (102), pp. 21-25

APPLIED METHODS OF STATISTICAL ANALYSIS.
STATISTICAL COMPUTATION AND SIMULATION

Proceedings
of the international workshop
Novosibirsk, 18-20 September 2019

Научное издание

ПРИКЛАДНЫЕ МЕТОДЫ СТАТИСТИЧЕСКОГО АНАЛИЗА.
СТАТИСТИЧЕСКИЕ РАСЧЕТЫ И МОДЕЛИРОВАНИЕ

ТРУДЫ
Международного семинара
Новосибирск, 18-20 сентября 2019 г.

На английском языке

Подписано в печать 16.08.2019. Формат 60×84 1/8.
Усл. печ. л. 71,75. Уч.-изд. л. 133,46.
Тираж 60 экз. Заказ № 1172.

Отпечатано в типографии
Новосибирского государственного технического университета
630073, г. Новосибирск, пр. К. Маркса, 20